ウィルクス

数理統計学

〈増訂新版〉

2

田中英之
岩本誠一　訳

東京図書株式会社

目　　次

第1巻の目次

第10章　線　形　推　定

10.1　序　　論

第8，9章では，標本論に関するいくつかの結果，すなわち与えられた c. d. f. $F(x)$ からの標本要素または成分からなる関数の確率論，について述べてきた．統計学の応用問題では，通常 $F(x)$ は未知であり，標本をとる主目的は，それを基礎として，$F(x)$ またはその性質の関係式や推論が導ける情報を得ることである．これらの関係式は標本要素（成分）の関数で述べられ，確率関係式で表現される．任意の標本を取り扱う前に，$F(x)$ について次の範囲の仮定が設けられる．つまり **2.2.1** で与えられた c. d. f. の基本特性のみを満たす $F(x)$ から，R_1 でのすべての x の値に対して $F(x)$ の値が完全に定まっているような $F(x)$ までである．一般に，与えられた任意の状態での $F(x)$ に関する仮定は，これら両端の間にあると考えられる．たとえば $F(x)$ が有限であるが未知の平均値 μ および分散 σ^2 を持つ場合，問題は $F(x)$ からの無作為標本要素の関数として，μ と σ^2 の推定量を定めることである．こうして8.2節では無限母集団からの標本の標本平均 $\bar{x}$ と標本分散 s^2 は μ と σ^2 に対する不偏推定量，すなわち $\mathscr{E}(\bar{x})=\mu$ かつ $\mathscr{E}(s^2)=\sigma^2$，であることを示した．さらに一般的にいえば，問題は分布からの標本に含まれている情報を利用して，分布関数に関する仮定だけではなく，母集団分布関数についての推定を求めることである．

統計学の応用問題では，分布の平均値，分散，共分散，回帰係数などの母集団分布のパラメータに対する不偏推定量を標本から求めることがよくある．このようなパラメータの比較的簡単な推定量は（線形推定量として知られているが）標本要素の1次形式として定めることができ，母集団 c. d. f. に関しては，標本成分の1次および2次モーメントの有限性以外に強い仮定は含んでいない．もちろん推定量についてはこの他にもあるが，それ

らは第 11, 12 章で考える．一般に母集団パラメータ θ に対する推定量は，標本から定められる観測可能な確率変数である．すなわちこれは推定するパラメータの未知の真値のかわりに用いられる標本要素の既知関数である． パラメータ θ に対する推定量を求める場合，重要なことは標本から推定値を求めることである． したがって θ が実際に値 θ_0 を持つような母集団からとった標本で推定するとき，その分布が真値 θ_0 のまわりにできるかぎり集まるようにしたい．測度に関する基準で見たように，線形推定量の場合のこの集中度は推定量の分散である．本章では，線形推定と，それを求める基準とさらに重要な統計学の問題への応用に注目しよう．その他の推定量に関する考察は第 11, 12 章で取り扱う．

8.2 と 8.5 節で指摘したように，標本平均は（もちろん， 標本要素の 1 次関数だが）無限または有限母集団の母集団平均の 1 つの不偏推定量である．しかし標本平均はわれわれが定めることのできる母集団平均の可能な線形不偏推定量のうちのただ 1 つに過ぎない．たとえば，$(x_1, \cdots, x_n)$ が平均 μ を持つ母集団（有限ないしは無限）からとった標本であれば，$a_1, \cdots, a_n$ を $a_1 + \cdots + a_n = 1$ となる既知定数としたとき，$a_1x_1 + \cdots + a_nx_n$ は μ に対する 1 つの 不偏推定量となっていることは明らかである． この型の推定量を**線形不偏推定量**と呼ぶ．μ に対する 1 つの線形不偏推定量を選ぶとき，その選定基準を考えなければならない．一般的な基準としては，すべての線形不偏推定量のうち，可能なかぎり最小の分散を持つ線形推定量を選ぶことである．このような推定量は**最小分散線形推定量**と呼ばれる．

同様に，$(x_1, \cdots, x_n)$ が平均値 μ と分散 σ^2 を持つ有限ないしは無限母集団からの標本であれば，平均値が σ^2 であるような任意の 2 次形式 $\sum_{\xi,\eta} a_{\xi\eta}(x_\xi - \bar{x})(x_\eta - \bar{x})$ は σ^2 に対する **2 次不偏推定量**である． σ^2 に対する可能なかぎり最小の分散を持つ唯一の 2 次不偏推定量が存在するならば， それを σ^2 に対する**最小分散 2 次推定量**と呼ぶ．しかし，任意の 2 次推定量は 2 次の項の 1 次関数であり，本章で展開する線形推定論の多くが 2 次推定量にも適用できることを注意しておこう．

通常，最小分散線形推定量は 2 次推定量により偏りがなく推定できる分散を持つ．

本章では，線形推定論と，それの母集団分布の平均および分散の推定への応用を取り扱う．これらの問題での線形推定の過程で用いられる確率変数の分布に関しては，確率変数の共分散行列の要素と平均値の有限性以外にはなんの仮定も必要としない． 2 次推定量に対する重要な確率変数は標本成分の 2 次形式である．

便宜上 θ の不偏推定量を $\mathscr{E}^{-1}(\theta)$ で表わす．T が θ の 1 つの不偏推定量となる観測可

能な確率変数であれば，すなわち

$$\mathscr{E}(T) = \theta$$

ならば，次のように書ける．

$$\mathscr{E}^{-1}(\theta) = T.$$

線形推定量および 2 次推定量に関する基本的な定理を次に示す．

10.1.1　$(x_1, \cdots, x_k)$ を次のような平均値を持つ確率変数とする．

$$\mathscr{E}(x_i) = \sum_j a_{ij}\theta_j, \qquad i = 1, \cdots, k$$

ただし $\theta_1, \cdots, \theta_k$ は未知のパラメータ，$\|a_{ij}\|$ はその要素が既知で（すなわち，パラメータ $\theta_1, \cdots, \theta_k$ に依存しない）正則な行列である．このとき

$$\mathscr{E}^{-1}(\theta_i) = \sum_j a^{ij}x_j, \qquad i = 1, \cdots, k$$

はそれぞれ $\theta_1, \cdots, \theta_k$ の線形不偏推定量である．ただし $\|a^{ij}\| = \|a_{ij}\|^{-1}$.

$A\theta = \mathcal{E}(x)$

この証明は読者に残しておく．

$\theta = A^{-1}\mathcal{E}(x)$

$\mathcal{E}^{-1}(\theta) = A^{-1}\mathcal{E}^{-1}(\mathcal{E}(x)) = A^{-1}x$

10.2　無作為標本からの母集団平均と母集団分散に対する最小分散線形推定量

（a）　母集団平均に対する最小分散線形推定量

$(x_1, \cdots, x_n)$ を平均値 μ と分散 σ^2 を持つ分布からの標本とする． 8.2 節でみたように標本平均は μ に対する不偏推定量である．ここで，$\bar{x}$ が μ の最小分散線形推定量であることを示そう．$\mathscr{E}^{-1}(\mu)$ を μ の任意の線形不偏推定量とする．すなわち

(10.2.1)　$$\mathscr{E}^{-1}(\mu) = a_1x_1 + \cdots + a_nx_n$$

ただし

(10.2.2)　$$a_1 + \cdots + a_n = 1.$$

$\mathscr{E}^{-1}(\mu)$ の分散については，**3.6.1a** から

(10.2.3)　$$\sigma^2(\mathscr{E}^{-1}(\mu)) = (a_1^2 + \cdots + a_n^2)\sigma^2.$$

$\sigma^2(\mathscr{E}^{-1}(\mu))$ は次の場合にのみ最小値を持つことがわかる．

(10.2.4)　$$a_1 = \cdots = a_n = \frac{1}{n}.$$

しかも，a の値をこのように選ぶと，$\mathscr{E}^{-1}(\mu)$ は $\bar{x}$ と一致する．したがって

10.2.1　$(x_1, \cdots, x_n)$ が平均値 μ と分散 σ^2 を持つ分布から抽出した大きさ n の標本であれば，$\bar{x}$ は μ に対する最小分散線形推定量となる．

$(x_1, \cdots, x_n)$ が有限母集団からの標本の場合にも同様な結果が得られる．証明は読者に残しておく．

(b) 母集団分散に対する最小分散2次推定量

2次推定量の問題を取り扱うには，まず Halmos (1946) の不偏推定論に関する次の定理を述べておくと都合がよい．

10.2.2　$(x_1, \cdots, x_n)$ を c.d.f. $F(x)$ からの標本とし，$g(x_1, \cdots, x_n)$ を平均値 θ と分散 $\sigma^2 < +\infty$ を持つ任意の統計量とする．$(i_1, \cdots, i_n)$ は整数 $(1, \cdots, n)$ の $n!$ 個すべての置換からなる組のうちの i 番目のもの（適当に指数をつけた）とし，$g_i(x_1, \cdots, x_n) = g(x_{i_1}, \cdots, x_{i_n})$ とする．$\bar{g} = \frac{1}{n!}\sum_{i=1}^{n!} g_i(x_1, \cdots, x_n)$ とすれば，$\mathscr{E}(\bar{g}) = \theta$ となり，$\bar{g}$ の分散は $g(x_1, \cdots, x_n)$ が確率1で $x_1, \cdots, x_n$ に関して対称でなければ $g(x_1, \cdots, x_n)$ の分散よりも小さい．確率1で対称の場合には $\bar{g}$ は $g(x_1, \cdots, x_n)$ と一致する．

10.2.2 を確かめよう．$(x_1, \cdots, x_n)$ は $F(x)$ からの無作為標本であるから，$\mathscr{E}(g_i) = \theta$, $i = 1, \cdots, n!$ であることに注目しよう．したがって $\mathscr{E}(\bar{g}) = \theta$ である．さらに，$\sigma^2(g_i) = \sigma^2$, $i = 1, \cdots, n!$, また

$$\sigma^2(\bar{g}) = \frac{1}{n!}\sigma^2 + \left(\frac{1}{n!}\right)^2 \sum_{i \neq j} \operatorname{cov}(g_i, g_j). \tag{10.2.5}$$

しかし $\operatorname{cov}(g_i, g_j) \leqslant \sigma^2$ であるから

$$\sigma^2(\bar{g}) \leqslant \sigma^2. \tag{10.2.6}$$

ここで (10.2.6) の等式が成り立つのは，すべての $i \neq j$ に対して標本空間 R_n で確率1で $g_i \equiv a + bg_j$ となるとき，かつそのときのみである．ただし a, b はそれぞれ値0と1をとるような定数である．ゆえに

$$g_i(x_1, \cdots, x_n) \equiv g_j(x_1, \cdots, x_n) \tag{10.2.7}$$

が標本空間 R_n のすべての点 $(x_1, \cdots, x_n)$（確率0の集合はのぞいて）と $i \neq j = 1, \cdots, n!$ に対して成り立つ．この条件は，$g(x_1, \cdots, x_n)$ が $(x_1, \cdots, x_n)$ に関して対称であることを意味している．**10.2.2** の結論は以上からただちに出てくる．

次の **10.2.2** の系は，ある適当な条件の下では標本分散 s^2 が母集団分布の分散 σ^2 の最小分散2次推定量になることを示している．証明は **10.2.2** をそのまま拡張すれば良いので，読者に練習問題として残しておく．

10.2.2 a　$(x_1, \cdots, x_n)$ が平均 μ と分散 σ^2 および有限な3次と4次のモーメントを持つ c.d.f. $F(x)$ からの標本で，$\{a_{\xi\eta}\}$ が定数の任意の組で，

$$Q = \sum_{\xi, \eta=1}^{n} a_{\xi\eta}(x_\xi - \bar{x})(x_\eta - \bar{x})$$

が平均値 σ^2 を持つならば，Q が最小の分散を持つ $\{a_{\xi\eta}\}$ の値は $a_{\xi\xi} = 1/(n-1)$, $\xi = 1, \cdots, n$, $a_{\xi\eta} = 0$, $\xi \neq \eta$ のときである．このとき，Q は標本分散 s^2 になる．

10.2.2 において，$(x_1, \cdots, x_n)$ が $F(x)$ からとった無作為標本であるという仮定を，$(x_1, \cdots, x_n)$ がベクトル確率変数でその c.d.f. $F(x_1, \cdots, x_n)$ が $x_1, \cdots, x_n$ に関して対称であるという仮定と入れ換えても **10.2.2** が成り立つことに注意しよう．**10.2.2** のこの拡張で，**10.2.2 a** も次のように拡張される．すなわち，有限母集団からの標本の場合，標本分散 s^2 は母集団分散 σ^2 の最小分散不偏推定量になっている．すなわち

10.2.2 b　$(x_1, \cdots, x_n)$ が分散 σ^2 を持つ有限母集団からの標本で，$\{a_{\xi\eta}\}$ が定数の任意の組で

$$Q = \sum_{\xi, \eta=1}^{n} a_{\xi\eta}(x_\xi - \bar{x})(x_\eta - \bar{x})$$

が平均 σ^2 を持つならば，Q が最小分散を持つ $\{a_{\xi\eta}\}$ の値は **10.2.2 a** で与えられる．このとき，Q は標本分散 s^2 となっている．

(c) 正規分布における μ と σ^2 に対する区間推定量

さらに $(x_1, \cdots, x_n)$ が分布 $N(\mu, \sigma^2)$ からとった標本であると仮定すれば，μ と σ^2 に対するいわゆる区間推定量（特別な場合）が得られる．この場合は，特別な注意が必要である．**8.4.3** から $\sqrt{n}(\bar{x} - \mu)/s$ は "スチューデント" 分布 $S(n-1)$ を持つことがわかっている．これから次がいえる．

$$P\left(t_1 < \frac{(\bar{x} - \mu)\sqrt{n}}{s} < t_2\right) = \int_{t_1}^{t_2} f_{n-1}(t)\,dt = \gamma \tag{10.2.8}$$

ただし $f_{n-1}(t)$ は (7.8.4) で与えられ，t_1 と t_2 は (10.2.8) の積分が値 γ を持つように選ばれたものである．しかし

$$P\left(t_1 < \frac{(\bar{x}-\mu)\sqrt{n}}{s} < t_2\right) = \gamma$$

は次式と同値である．

(10.2.9)
$$P\left(\bar{x} - t_2\frac{s}{\sqrt{n}} < \mu < \bar{x} - t_1\frac{s}{\sqrt{n}}\right) = \gamma.$$

このように，$(\bar{x} - t_2(s/\sqrt{n}),\ \bar{x} - t_1(s/\sqrt{n}))$ は点 μ を含む確率が γ である観測可能なランダムな区間である．この区間を μ に対する $100\gamma\%$ **信頼区間**と呼び，γ を**信頼係数**と呼ぶ．これは，パラメータの"真の"値を含む特定の確率を持つ（実現可能あるいは観測可能な）確率区間を用いることによってパラメータの1つの推定量を設定する1方法である．このような区間を**区間推定量**と呼ぶ．さらに一般的な条件での区間推定法の議論は，第11，12章で行なう．

上の例では，区間の長さは $(t_2 - t_1)s/\sqrt{n}$ である．これは，γ が与えられると t_1 と t_2 が（与えられた γ に対して）$t_2 = -t_1 = t_{n-1,\gamma}$ で，$t_{n-1,\gamma}$ が

(10.2.10)
$$\int_{-t_{n-1,\gamma}}^{+t_{n-1,\gamma}} f_{n-1}(t)\,dt = \gamma$$

を満たすように選ばれたとき，与えられた s に対して最小となる．この場合には μ に対する $100\gamma\%$ 信頼区間は $\bar{x}$ を中心とする次のような区間である．

(10.2.11)
$$\bar{x} \pm t_{n-1,\gamma}\frac{s}{\sqrt{n}}.$$

同様に，$(x_1, \cdots, x_n)$ が分布 $N(\mu, \sigma^2)$ を持つ母集団からの標本であれば，$(n-1)s^2/\sigma^2$ がカイ2乗分布 $C(n-1)$ を持つ［**8.4.2** を見よ］ことから，σ^2 に対する区間推定量を設定することができる．なぜなら

(10.2.12)
$$P\left(\chi_1^2 < \frac{(n-1)s^2}{\sigma^2} < \chi_2^2\right) = \int_{\chi_1^2}^{\chi_2^2} dF_{n-1}(\chi^2) = \gamma$$

であるからである．ただし，$dF_{n-1}(\chi^2)$ は (7.8.1) で与えられている．ゆえに (10.2.12) の積分の値は γ になる．しかし (10.2.12) は次式と同値である．

$$P\left(\frac{(n-1)s^2}{\chi_2^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_1^2}\right) = \gamma.$$

このことから $((n-1)s^2/\chi_2^2,\ (n-1)s^2/\chi_1^2)$ が σ^2 に対する $100\gamma\%$ 信頼区間であることは明らかである．もちろん，(10.2.12) を満足する χ_1^2 と χ_2^2 の選び方には多数の方法がある．一般には下のように選ぶ．

(10.2.13)
$$\int_0^{\chi_1^2} dF_{n-1}(\chi^2) = \int_{\chi_2^2}^{\infty} dF_{n-1}(\chi^2) = \frac{1-\gamma}{2}.$$

しかしながら，長さの平均値が最小になる信頼区間は $(1/\chi_1^2 - 1/\chi_2^2)$ が (10.2.12) を満たすように最小化されたときに得られる．

10.3　線形回帰分析におけるパラメータの推定量

(a)　回帰係数に対する推定量

ここでは，回帰分析の推定問題，実験計画法およびその関連問題に必要となる **10.2.1** の一般化を考えてみよう．$y_1, \cdots, y_n$，$n > k$ を n 個の独立な確率変数とする．ただし分散はすべて σ^2 に等しいが，平均は次の回帰関数で与えられるものとしよう．

(10.3.1)
$$\mathscr{E}(y_\xi) = \beta_1 x_{1\xi} + \cdots + \beta_k x_{k\xi}, \qquad \xi = 1, \cdots, n.$$

ただし $(x_{1\xi}, \cdots, x_{k\xi})$，$\xi = 1, \cdots, n$ は既知の（実）ベクトルであるが，$\beta_1, \cdots, \beta_k$ は未知の（実数）パラメータである．これは推定すべきもので**回帰係数**という．パラメータ σ^2 は，通常未知で，**残差分散**といわれる．ここで $x_1, \cdots, x_k$ を導入し，これらを確率変数に対応させて，**固定変数**と呼ぼう．この場合，$(x_{1\xi}, \cdots, x_{k\xi})$，$\xi = 1, \cdots, n$ は，固定変数の n 個の特定の値の組である．習慣上，$x_{1\xi} = 1$，$\xi = 1, \cdots, n$ とするが，いまのところはこうしない方が便利であろう．$\beta_1, \cdots, \beta_k$ の最小分散推定量がある適当な条件のもとで存在し，σ^2 に対しては比較的簡単な不偏推定量が存在することを示そう．

まず $\beta_1, \cdots, \beta_k$ の推定を考える．$\mathscr{E}^{-1}(\beta_i)$ を β_i の任意の不偏線形推定量とする．すなわち

(10.3.2)
$$\mathscr{E}^{-1}(\beta_i) = \sum_\xi c_{i\xi} y_\xi, \qquad i = 1, \cdots, k.$$

このとき

(10.3.3)
$$\mathscr{E}(\mathscr{E}^{-1}(\beta_i)) = \sum_\xi \sum_j \beta_j c_{i\xi} x_{j\xi} \equiv \beta_i, \qquad i = 1, \cdots, k.$$

これから明らかに $c_{i\xi}$ は次式を満足しなければならない．

(10.3.4)
$$\sum_\xi c_{i\xi} x_{j\xi} = \delta_{ij}.$$

ただし δ_{ij} はクロネッカーの δ である．$\mathscr{E}^{-1}(\beta_i)$ の分散は

(10.3.5)
$$\sigma^2(\mathscr{E}^{-1}(\beta_i)) = \sum_\xi c_{i\xi}^2 \sigma^2$$

で与えられる．(10.3.4) の条件の下で $c_{i\xi}$ に関して $\sigma^2(\mathscr{E}^{-1}(\beta_i))$ を最小にするには，$c_{i\xi}$ は次のような形をしていなければならない．

(10.3.6)
$$c_{i\xi} = \sum_j \lambda_{ij} x_{j\xi}, \qquad i = 1, \cdots, k, \qquad \xi = 1, \cdots, n.$$

ここで λ_{ij} は次の等式を満足しなければならない．

(10.3.7)
$$\sum_{j'} \lambda_{ij'} a_{j'j} = \delta_{ij}$$

ただし

(10.3.8)
$$a_{j'j} = \sum_\xi x_{j'\xi} x_{j\xi}.$$

行列 $\|a_{j'j}\|$ が正則 ($(x_{i1}, \cdots, x_{in})$, $i = 1, \cdots, k$ が1次独立なベクトルのときかつそのときにのみ) ならば，(10.3.7) の解が次式になることは明白である．

(10.3.9)
$$\|\lambda_{ij}\| = \|a_{ij}\|^{-1} = \|a^{ij}\|.$$

したがって，β_i に対する最小分散線形推定量を b_i とすれば

(10.3.10)
$$b_i = \sum_j a^{ij} a_{j0}$$

となる．ただし a_{j0}, $j = 1, \cdots, k$ は次で定義される確率変数である．

(10.3.11)
$$a_{j0} = \sum_\xi x_{j\xi} y_\xi = a_{0j}.$$

b_i の分散を求めるために，(10.3.9) の λ_{ij} を (10.3.6) に代入し，次にその $c_{i\xi}$ を (10.3.5) に代入すると次式が得られる．

(10.3.12)
$$\sigma^2(b_i) = \sum_{j, j'=1}^{k} a^{ij} a^{ij'} a_{j'j} \sigma^2 = \sum_j a^{ij} \delta_{ij} \sigma^2 = a^{ii} \sigma^2.$$

同様にして，b_i, $i = 1, \cdots, k$ の共分散行列が次のようになることがわかる．

(10.3.13)
$$\|a^{ij} \sigma^2\|.$$

以上をまとめると，次の Markov (1900) の定理が得られる．

10.3.1　y_ξ, $\xi = 1, \cdots, n$ を平均 $\sum_{i=1}^{k} \beta_i x_{i\xi}$, $\xi = 1, \cdots, n \geqslant k$ と，すべて σ^2 に等しい分散を持つ独立な確率変数とする．ただし $(x_{i1}, \cdots, x_{in})$, $i = 1, \cdots, k$ は既知の1次独立なベクトルである．このとき回帰係数 β_i の最小分散線形推定量は，(10.3.10) で定義した b_i, $i = 1, \cdots, k$ となる．また b_i の共分散行列は $\|a^{ij}\sigma^2\|$ である．ただし $a_{ij} = \sum_\xi x_{i\xi} x_{j\xi}$, $i, j = 1, \cdots, k$ かつ $\|a^{ij}\| = \|a_{ij}\|^{-1}$ である．

ここで，もし次の2乗和を $\beta_1, \cdots, \beta_k$ に関して最小にするならば，

(10.3.14)
$$Q = \sum_\xi (y_\xi - \beta_1 x_{1\xi} - \cdots - \beta_k x_{k\xi})^2,$$

$\|a_{ij}\|$ が正則なときは，β_i に対する**最小2乗推定量**，すなわち (10.3.14) を最小にするような β_i の値は b_i となることに注意しよう．詳細については読者に残しておく．このようにして，回帰係数の線形推定量に関する次の定理が得られる．

10.3.2　**10.3.1** の条件の下では，回帰係数 β_i の最小分散線形推定量は β_i の最小2乗推定量に等しい．

β_i の推定量に対する最小分散からの接近は Markov (1900) によるが，他方ごく初期の最小2乗法に関しては，Gauss (1809 *a*) によるものである．**10.3.1** と **10.3.2** を結び合わせて，通常，**ガウス=マルコフの定理**と呼ぶ．

多くの回帰問題では，$x_{1\xi} = 1$, $\xi = 1, \cdots, n$, すなわち $y_\xi = \beta_1 + \beta_2 x_{2\xi} + \cdots + \beta_k x_{k\xi}$ と仮定している．この場合には，**10.3.2** を用いると，β_i の最小分散線形推定量 b_i は次のように表わされる．

(10.3.15)
$$b_1 = \bar{y} - b_2 \bar{x}_2 - \cdots - b_k \bar{x}_k$$
$$b_{i'} = \sum_{j'=2}^{k} A^{i'j'} A_{j'0}, \qquad i' = 2, \cdots, k$$

ただし

$$\bar{y} = \frac{1}{n} \sum_\xi y_\xi, \qquad \bar{x}_{i'} = \frac{1}{n} \sum_\xi x_{i'\xi}, \qquad i' = 2, \cdots, k.$$

(10.3.16)
$$A_{i'j'} = \sum_\xi (x_{i'\xi} - \bar{x}_{i'})(x_{j'\xi} - \bar{x}_{j'}), \qquad i', j' = 2, \cdots, k$$
$$A_{j'0} = \sum_\xi (\bar{x}_{j'\xi} - \bar{x}_{j'})(y_\xi - \bar{y})$$

そして

(10.3.17)
$$\|A^{i'j'}\| = \|A_{i'j'}\|^{-1}.$$

さらに，a_{1i} で $x_{1\xi} = 1$, $\xi = 1, \cdots, n$ と置けば，(10.3.13) より $b_{i'}$, $i' = 2, \cdots, k$ の共分散行列は次のようになることがわかる．

(10.3.18)
$$\|A^{i'j'}\sigma^2\|,\qquad i', j' = 2, \cdots, k$$

しかるに

(10.3.19)
$$\sigma^2(b_1) = \left[\frac{1}{n} + \sum_{i', j'=2}^{k} A^{i'j'}\bar{x}_{i'}\bar{x}_{j'}\right]\sigma^2$$
$$\sigma(b_1, b_{i'}) = \left[\sum_{j'=2}^{k} A^{i'j'}\bar{x}_{j'}\right]\sigma^2.$$

ガウス=マルコフの定理と秤量問題に関する注意　(10.3.1) において，$\beta_1, \cdots, \beta_k$ がそれぞれ $o_1, \cdots, o_k$ である k 個の物体の“真の”重みを表わしているとしよう．ここで，これらの物体の種々の組合せを化学秤（左右の秤皿を持った秤）に乗せた場合を考えよう．o_i が左の皿に乗っているとき $x_{i\xi} = +1$，o_i が右の皿に乗っているとき $x_{i\xi} = -1$，そして o_i を計っていないときは $x_{i\xi} = 0$，$i = 1, \cdots, k$ とする．このとき $\beta_1 x_{1\xi} + \cdots + \beta_k x_{k\xi}$ は構成が $x_{1\xi}, \cdots, x_{k\xi}$ によって決まる物体の組の第 ξ 回目の測定の“真の”表示である．n 個の異なった測り方に対して，測定割当行列 $\|x_{i\xi}\|$，$i = 1, \cdots, k$，$\xi = 1, \cdots, n$ は -1 と 0 と $+1$ だけからなっている．

もし，秤が正しければ，ξ 番目の測定で“実際の”値 y_ξ は，平均値が $\beta_1 x_{1\xi} + \cdots + \beta_k x_{k\xi}$ になるような確率変数と考えてもよい．$y_1, \cdots, y_n$ を n 回の測定で得られる n 個の確率変数とし，これらの確率変数が等分散 σ^2 を持つ独立変数とすれば，物体 $o_1, \cdots, o_k$ の“真の”重み $\beta_1, \cdots, \beta_k$ の最小分散推定量 $b_1, \cdots, b_k$ は (10.3.10) で与えられ，推定量の共分散行列は (10.3.13) で与えられる．k はすべての β の推定量を与える測定割当行列を求めることのできる n の最小値である．$(\beta_1, \cdots, \beta_k)$ のベクトル推定量 $(b_1, \cdots, b_k)$ を種々の意味で“最良に”与える測定割当行列を構成する問題は，Hotelling (1944), Kishen (1945), Mood (1946) などによって広く研究されている．

(b)　残差分散 σ^2 の推定量

ここでは，残差分散 σ^2 の2次不偏推定量を構成する問題を考えよう．

(10.3.20)
$$z_\xi = y_\xi - \sum_{i=1}^{k} \beta_i x_{i\xi}$$

(10.3.21)
$$\tilde{y}_\xi = \sum_{i=1}^{k} b_i x_{i\xi},$$

(10.3.22)
$$S = \sum_{\xi=1}^{n} z_\xi^2,\qquad S_1 = \sum_{\xi=1}^{n} (y_\xi - \tilde{y}_\xi)^2,\qquad S_2 = \sum_{i,j=1}^{k} a_{ij}(b_i - \beta_i)(b_j - \beta_j)$$

と置くと

(10.3.23)
$$S = S_1 + S_2$$

となるが，

(10.3.24)
$$\mathscr{E}(S) = n\sigma^2$$

であって，共分散行列 (10.3.13) を用いると次式が得られる．

(10.3.25)
$$\mathscr{E}(S_2) = \sum_{i,j=1}^{k} a^{ij}a_{ij}\sigma^2 = k\sigma^2.$$

$\mathscr{E}(S) = \mathscr{E}(S_1) + \mathscr{E}(S_2)$ であるから

(10.3.26)
$$\mathscr{E}(S_1) = (n - k)\sigma^2.$$

S_1 はパラメータ $\beta_1, \cdots, \beta_k$，σ^2 に無関係だから，観測可能であり，次の結果が得られる．

10.3.3　**10.3.1** の条件の下で，$S_1/(n-k)$ は σ^2 に対する不偏推定量である．

S, S_1, S_2 の平均値はそれぞれ $n\sigma^2$, $(n-k)\sigma^2$, $k\sigma^2$ であることに注意しよう．n, $n-k$, k は S, S_1, S_2 の自由度を表わしている．

(10.3.22) における S_1 を2乗して ξ に関する和をとれば，

(10.3.27)
$$S_1 = a_{00} - \sum_{i,j} a_{ij}b_i b_j = a_{00} - \sum_{i,j} a^{ij}a_{i0}a_{j0}$$

となる．すなわち，S_1 は次のように簡単な形に書くことができる．

(10.3.27 a)
$$S_1 = \frac{|a_{i0j0}|}{|a_{ij}|}.$$

ただし，$a_{00} = \sum_\xi y_\xi^2$ で

(10.3.28)
$$|a_{i0j0}| = \begin{vmatrix} a_{00} & a_{01} & \cdots & a_{0k} \\ a_{10} & a_{11} & \cdots & a_{1k} \\ \cdot & & \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & & \cdot\ \cdot \\ a_{k0} & a_{k1} & \cdots & a_{kk} \end{vmatrix}.$$

(10.3.27) と (10.3.27 a) が同値であることは，(10.3.28) の行列を第1行第1列で境界展開してみれば明らかである [たとえば，Bôcher (1907) を見よ]．

回帰係数のいくつか，たとえば $\beta_{k_1+1}, \beta_{k_1+2}, \cdots, \beta_k$ の値が既知であれば，上で y_ξ を y'_ξ で置き換え（ただし $y'_\xi = (y_\xi - \beta_{k_1+1}x_{k_1+1\xi} - \cdots - \beta_k x_{k\xi})$），$k$ を k_1 で置き換えて，前述と同様の解析を行なうことができる．このようにして，容易に**10.3.1**，**10.3.2**，**10.3.3** の変形が得られる．

(c) 正規回帰論における回帰推定量の分布

さらに確率変数 y_ξ, $\xi=1,\cdots,n$ が独立で分布 $N(\beta_1 x_{1\xi}+\cdots+\beta_k x_{k\xi},\sigma^2)$ を持つと仮定すれば，統計学の応用上重要な結果が得られる．z_ξ を (10.3.20) で定義すれば，確率変数 z_ξ, $\xi=1,\cdots,n$ の p.e. は

$$dF(z_1,\cdots,z_n)=(1/\sqrt{2\pi}\sigma)^n\exp\left(-\frac{1}{2\sigma^2}\sum_\xi z_\xi^2\right)dz_1\cdots dz_n. \tag{10.3.29}$$

(10.3.22) より S_1 および S_2 は

$$\begin{aligned} S_1&=\sum_\xi\left[z_\xi-\sum_{i,j=1}^k a^{ij}\left(\sum_{\eta=1}^n x_{i\eta}z_\eta\right)x_{i\xi}\right]^2\\ S_2&=\sum_{\xi,\eta}\sum_{i,j}a^{ij}x_{i\xi}x_{j\eta}z_\xi z_\eta \end{aligned} \tag{10.3.30}$$

となり，これはそれぞれ階数 $n-k$, k の行列を持つ z_ξ, $\xi=1,\cdots,n$ に関する2次形式である．S が階数 n の行列を持つ z_ξ, $\xi=1,\cdots,n$ に関する2次形式であるから，コクランの定理 **8.4.4** より，S_1/σ^2 と S_2/σ^2 はそれぞれカイ2乗分布 $C(n-k)$, $C(k)$ に従って独立に分布する．さらに，(10.3.10), (10.3.11) から，b_i は y_ξ の1次関数である．すなわち

$$b_i=\sum_\xi\sum_j a^{ij}x_{j\xi}y_\xi,\qquad i=1,\cdots,k \tag{10.3.31}$$

で，平均 β_i と (10.3.13) で与えられる共分散行列を持つ．したがって，**7.4.4** により，$\|a_{ij}\|$ が正則ならば，b_i は k 次元正規分布 $N(\{\beta_i\},\|a^{ij}\sigma^2\|)$ を持つ．実際，正規性の仮定の下では S_1, $(b_1,\cdots,b_k)$ は，確率変数の2つの独立な組になる．これは $(S_1,b_1,\cdots,b_k)$ の特性関数 $\mathscr{E}(e^{\mathrm{i}tS_1+\mathrm{i}[t_1b_1+\cdots+t_kb_k]})$ を求めれば示される．すなわち

$$(1-2\sigma^2\mathrm{i}t)^{-\frac{1}{2}(n-k)}\cdot\exp\left\{-\frac{\sigma^2}{2}\sum_{i,j}a^{ij}t_it_j+\mathrm{i}\sum_i\beta_it_i\right\} \tag{10.3.32}$$

となる．この結果，次の正規回帰論における重要な定理が得られる．

10.3.4　確率変数 y_ξ, $\xi=1,\cdots,n$ が独立で分布 $N(\beta_1x_{1\xi}+\cdots+\beta_kx_{k\xi},\sigma^2)$, $\xi=1,\cdots,n$ を持つならば，次が成り立つ．ただし (10.3.8) で定義された行列 $\|a_{ij}\|$, $i,j=1,\cdots,k$ は正則とする．

(i) (10.3.10) で定義された b_i は回帰係数 β_i, $i=1,\cdots,k$ の線形不偏推定量であり，k 次元正規分布 $N(\{\beta_i\},\|a^{ij}\sigma^2\|)$ を持つ．

(ii) S_1 と $(b_1,\cdots,b_k)$ は確率変数の独立な組である．

(iii) S_1/σ^2 と S_2/σ^2 は独立でそれぞれカイ2乗分布 $C(n-k)$, $C(k)$ を持つ．

10.4 正規回帰論におけるパラメータに対する区間推定量と楕円体推定量

10.3.4 の条件の下では，S_1 と $(b_1,\cdots,b_k)$ は独立な分布を持つ確率変数の2つの組であるから，S_1 と任意に与えられた b_i が独立であることは明らかである．しかし，b_i は分布 $N(\beta_i,a^{ii}\sigma^2)$ を持ち，S_1/σ^2 はカイ2乗分布 $C(n-k)$ を持つので，**7.8.3** から任意の b_i に対して

$$\frac{(b_i-\beta_i)\sqrt{n-k}}{\sqrt{a^{ii}S_1}} \tag{10.4.1}$$

は“スチューデント”分布 $S(n-k)$ を持つことがわかる．また，同様に信頼区間 (10.2.11) は

$$b_i\pm(t_{n-k,\gamma})\sqrt{\frac{a^{ii}S_1}{(n-k)}} \tag{10.4.2}$$

となり，これは β_i に対する $100\gamma\%$ 信頼区間となっている．ただし，$t_{n-k,\gamma}$ は $n-1$ を $n-k$ で置き換えた (10.2.10) を満足するものである．

$i=1,\cdots,k$ についての信頼区間 (10.4.2) を構成する場合，すべての β を被う信頼区間の個数の平均値が γk であることは明らかである．しかし，b が独立でないならば（ただし独立となるのは $\|a^{ij}\|$ が対角行列の場合に限る），これは各 β を同時に被う信頼区間の確率にはならない．そこで，パラメータの点 $(\beta_1,\cdots,\beta_k)$ を被う，確率が γ となる k 次元 β 空間での簡単な確率領域 R_γ を設定することができるかどうかという問題が起ってくる．この例では，S_1/σ^2 と S_2/σ^2 がそれぞれカイ2乗分布 $C(n-k)$, $C(k)$ に従う独立な分布であることから，容易にこの領域を求めることができる．なぜなら，**7.8.5** から

$$\frac{(n-k)S_2}{kS_1} \tag{10.4.3}$$

がスネディッカー分布 $S(k,n-k)$ を持つことは明白である．S_2 に (10.3.22) の S_2 を

代入すれば，

(10.4.4)
$$P\left(\left(\frac{n-k}{kS_1}\right)\sum_{i,j=1}^{k}a_{ij}(b_i-\beta_i)(b_j-\beta_j)<\mathscr{F}_{k,n-k,\gamma}\right)$$
$$=\int_0^{\mathscr{F}_{k,n-k,\gamma}}dF_{k,n-k}(\mathscr{F})=\gamma$$

となる．ただし $dF_{k,n-k}(\mathscr{F})$ はスネディッカー分布 $S(k,n-k)$ の p.e. であり，その一般形は (7.8.9) で与えられる．しかし (10.4.4) は次のように述べることができる．

(10.4.5)
$$P((\beta_1,\cdots,\beta_k)\in R_\gamma)=\gamma$$

ここで R_γ は $(b_1,\cdots,b_k)$ を中心に持つ次式で表わされる β 空間のランダムな楕円体の内部である．

(10.4.6)
$$\sum_{i,j=1}^{k}a_{ij}(\beta_i-b_i)(\beta_j-b_j)=\frac{kS_1\mathscr{F}_{k,n-k,\gamma}}{n-k}.$$

この楕円体はパラメータ点 $(\beta_1,\cdots,\beta_k)$ に対する 100γ% 信頼領域の1例である．楕円体 (10.4.6) は，$(\beta_1,\cdots,\beta_k)$ に対する領域推定量と呼ばれる．同様に，β の任意の部分集合に対する領域推定量を設定することができる．これは練習問題として読者に残しておく．単独の β，たとえば β_i に対する信頼区間推定量 (10.4.2) は区間（1次元領域）推定量である．

信頼楕円体の概念は，第18章で述べる Hotelling (1931) による一般化された"スチューデント"分布とともに導入された．

10.5　同時信頼区間：多次元の場合

前節では，標本回帰係数ベクトル $(b_1,\cdots,b_k)$ が母集団回帰係数ベクトル $(\beta_1,\cdots,\beta_k)$ に対する信頼楕円体を求めるのにいかに利用されるかを見てきた．特に分散分析の問題では，後節で述べるように，既知または観測可能な共分散行列を持つ k 次元正規確率変数の成分の1次結合（係数は既知とする）の多くに対して同時に成り立つ信頼区間を構成することが望ましい．この問題は Duncan (1952)，Dwass (1959)，Roy と Bose (1953)，Roy (1954)，Scheffé (1953)，Tukey (1953) などによって考えられてきた．

（a） 同時信頼区間に対する確率不等式

まず，パラメータが h 個のそれぞれの信頼区間に同時に含まれるような確率に対する不等式を考えてみよう．$\mu_1,\cdots,\mu_h$ を未知のパラメータとし，$(\underline{\mu}_1,\overline{\mu}_1),\cdots,(\underline{\mu}_h,\overline{\mu}_h)$ を信頼係数 $1-(1-\gamma)/h$ を持つ任意の区間とする．E_i を $(\underline{\mu}_i,\overline{\mu}_i)$ が μ_i を含む事象とし，$\overline{E}_i$ をその余事象とする，ただし $i=1,\cdots,h$．このとき

$$P(\overline{E}_i)=\frac{1}{h}(1-\gamma),\qquad i=1,\cdots,h$$

となる．

すべての事象 $E_1,\cdots,E_h$ が同時に起る確率は $P(E_1\cap\cdots\cap E_h)$ である．しかし

(10.5.1)
$$P(E_1\cap\cdots\cap E_h)=1-P(\overline{E}_1\cup\cdots\cup\overline{E}_h)$$

また

(10.5.2)
$$P(\overline{E}_1\cup\cdots\cup\overline{E}_h)\leqslant P(\overline{E}_1)+\cdots+P(\overline{E}_h).$$

したがって

(10.5.3)
$$P(E_1\cap\cdots\cap E_h)\geqslant 1-[P(\overline{E}_1)+\cdots+P(\overline{E}_h)].$$

すなわち

(10.5.4)
$$P(E_1\cap\cdots\cap E_h)\geqslant\gamma.$$

まとめると次のようになる．これは Tukey (1953) による．

10.5.1　$\mu_1,\cdots,\mu_h$ は未知のパラメータで，$(\underline{\mu}_1,\overline{\mu}_1),\cdots,(\underline{\mu}_h,\overline{\mu}_h)$ をそれぞれ $\mu_1,\cdots,\mu_h$ に対する $100\left[1-\frac{1}{h}(1-\gamma)\right]\%$ 信頼区間とする．このとき，これらの信頼区間がそれぞれ $\mu_1,\cdots,\mu_h$ を同時に含む確率は少なくとも γ である．

さてここで，特に後節で述べるモデルⅠ分散分析問題に適用できるいくつかの特殊な条件の下で，多くの信頼区間を同時に満たす問題を考えよう．

（b） シェフィーの方法

Scheffé (1953) の方法を述べよう．

10.5.2　$(u_1,\cdots,u_k)$ を次の分布を持つ k 次元確率変数とする．

$$N(\{\mu_i\},\|\sigma^2a_{ij}\|),\qquad i,j=1,\cdots,k.$$

ただし $\|a_{ij}\|$ は正則で，対称な既知行列とし，σ^2 は未知とする．v/σ^2 を $(u_1,\cdots,u_k)$ と独立な確率変数とし，自由度 m のカイ2乗分布を持つとす

る．また，$\mathscr{F}_{k,m,\gamma}$ はスネディッカー分布 $S(k,m)$ の $100\gamma\%$ 点とし，$\delta=\sqrt{kv\mathscr{F}_{k,m,\gamma}/m}$ とする．$\mathscr{C}$ を $c_1,\cdots,c_k$ がすべては0でないような実ベクトル $(c_1,\cdots,c_k)$ の全体の集合とすれば，次の不等式が $\mathscr{C}$ におけるすべての $(c_1,\cdots,c_k)$ に対して同時に成立する確率は γ である．

$$\sum_i c_iu_i-\delta\sqrt{\sum_{i,j}a_{ij}c_ic_j}\leqslant\sum_i c_i\mu_i\leqslant\sum_i c_iu_i+\delta\sqrt{\sum_{i,j}a_{ij}c_ic_j}. \tag{10.5.5}$$

$c_1,\cdots,c_k$ がすべて0である場合には（これは $\mathscr{C}$ には含まれないが），(10.5.5) は確率1で成り立つ．

10.5.2 の証明は，$\dfrac{1}{\sigma^2}\sum_{i,j}a^{ij}(u_i-\mu_i)(u_j-\mu_j)$ と v/σ^2 がおのおの自由度 k, m のカイ2乗分布を持つ独立な確率変数であることに注目すればよい．すなわち

$$\frac{m}{kv}\sum_{i,j}a^{ij}(u_i-\mu_i)(u_j-\mu_j)$$

はスネディッカー分布 $S(k,m)$ を持つ．ゆえに，$\mathscr{F}_{k,m,\gamma}$ がこの分布の $100\gamma\%$ 点であれば

$$P\left(\sum_{i,j}a^{ij}(u_i-\mu_i)(u_j-\mu_j)<\delta^2\right)=\gamma \tag{10.5.6}$$

となる．ただし

$$\delta^2=\frac{kv}{m}\mathscr{F}_{k,m,\gamma}.$$

これからは便宜上，k 次元の幾何学の概念やその用語を用いよう．

$$\sum_{i,j}a^{ij}(\mu_i-u_i)(\mu_j-u_j)<\delta^2 \tag{10.5.7}$$

となるような $(\mu_1,\cdots,\mu_k)$ 空間の点集合は，$(u_1,\cdots,u_k)$ に中心を持つ真のパラメータ点 $(\mu_1,\cdots,\mu_k)$ に対する $100\gamma\%$ 信頼楕円体の内部である．$(\mu_1,\cdots,\mu_k)$ の空間において，この楕円体に接するすべての可能な，$(k-1)$ 次元超平面の平行な組の間に含まれる点集合を考えれば，この点集合は楕円体の内部 (10.5.7) を構成する．もちろんこの集合に含まれる確率は γ である．ここでは次のことを示さなくてはならない．すなわち，$\mathscr{C}$ における任意の $(c_1,\cdots,c_k)$ の選び方に対して，$(\mu_1,\cdots,\mu_k)$ 空間の2つの平行な $(k-1)$ 次元超平面

$$\sum_i c_i\mu_i=\sum_i c_iu_i\pm\delta\sqrt{\sum_{i,j}a_{ij}c_ic_j} \tag{10.5.8}$$

が楕円体

$$\sum_{i,j}a^{ij}(\mu_i-u_i)(\mu_j-u_j)=\delta^2 \tag{10.5.9}$$

に接することである．2つの超平面 (10.5.8) の間の任意の点 $(\mu_1,\cdots,\mu_k)$ が (10.5.5) を満足することは明らかである．いま，$\mu_i-u_i=y_i$ と置くと，(10.5.9) は

$$\sum_{i,j}a^{ij}y_iy_j=\delta^2 \tag{10.5.10}$$

となり，$(y_1,\cdots,y_k)$ の空間での任意の超平面は次式で表わされる．

$$\sum_i c_iy_i=d. \tag{10.5.11}$$

ここで，超平面 (10.5.11) が楕円体 (10.5.10) に接するような d の2つの値を見つけなければならない．ラグランジュ乗数 λ を用いて

$$\Phi=\frac{1}{2}\lambda\left(\delta^2-\sum_{i,j}a^{ij}y_iy_j\right)+\sum_i c_iy_i \tag{10.5.12}$$

の $(y_1,\cdots,y_k)$ 空間における停留点を求める．y_j に関して微分すると

$$-\lambda\sum_i a^{ij}y_i+c_j=0,$$

$$y_i=\frac{1}{\lambda}\sum_j a_{ij}c_j. \tag{10.5.13}$$

これを (10.5.10) に代入すると

$$\lambda=\pm\frac{1}{\delta}\sqrt{\sum_{i,j}a_{ij}c_ic_j} \tag{10.5.14}$$

(10.5.14), (10.5.13), (10.5.11) から，

$$d=\pm\delta\sqrt{\sum_{i,j}a_{ij}c_ic_j}.$$

この d の値を (10.5.11) に代入し，$y_i=\mu_i-u_i$ を用いれば，指定された $(c_1,\cdots,c_k)$ に対する2つの接超平面の式として (10.5.8) が得られる．

最後に $\mathscr{C}$ におけるベクトルとして，有限な N 個だけをとろう．これら N 個のベクトルに対応する超平面の平行な対すべての N 個の組の間にあるような $(\mu_1,\cdots,\mu_k)$ 空間における点集合は楕円体 (10.5.9) に外接するランダムな k 次元多面体 G_k である．G_k は楕円体を含むから，G_k に含まれる確率は楕円体に含まれる確率 γ を越える．したがって，$\mathscr{C}$ における任意の有限な N 個のベクトルに対応する N 個の不等式が同時に成り立

つ確率は γ より大きい．たとえば，$c_i=1$，$c_j=-1$，$i>j=1,\cdots,k$ でその他の c はすべて0に等しいとすれば，$\mathscr{C}$ において $N=k(k-1)/2$ 個のベクトルからなる部分集合が得られる．すると次の $k(k-1)/2$ 個の不等式がすべて同時に成り立つ確率は γ を超える．

(10.5.5 a)
$$(u_i-u_j)-\delta\sqrt{a_{ii}+a_{jj}-2a_{ij}}\leqslant(\mu_i-\mu_j)$$
$$\leqslant(u_i-u_j)+\delta\sqrt{a_{ii}+a_{jj}-2a_{ij}},\qquad i>j=1,\cdots,k.$$

(c)　トゥキーの方法

可能なすべての差 $\mu_i-\mu_j$，$i>j=1,\cdots,k$ の場合に対しても，(10.5.5 a) と同じような結果が得られるが，Tukey (1953) はこれを σ に対する“スチューデント化された範囲”の推定量を用いて求めている．$(z_1,\cdots,z_k)$ を $N(\mu,d^2\sigma^2)$ からの大きさ k の標本とする．ここで d^2 は既知の正の定数，v/σ^2 は $(z_1,\cdots,z_k)$ と独立な確率変数で，自由度 m のカイ2乗分布を持つとする．R を $(z_1,\cdots,z_k)$ の範囲とする．すなわち $R=\max(z_1,\cdots,z_k)-\min(z_1,\cdots,z_k)$．このとき確率変数 $\sqrt{m}\,R/\sqrt{v}\,d$ をスチューデント化された範囲 $R_{k,m}$ と呼ぶ．与えられた信頼係数 γ に対して，$R_{k,m,\gamma}$ を次のように定義された $R_{k,m}$ の分布の右側 $100\gamma\%$ 点とする．

(10.5.15)
$$P(R_{k,m}\leqslant R_{k,m,\gamma})=\gamma.$$

Tukey (1953) の結果は

10.5.3　$(u_1,\cdots,u_k)$ を分布 $N(\{\mu_i\},\|\sigma^2a_{ij}\|)$ を持つ k 次元確率変数とする．ただし $a_{ii}=a^2$，$a_{ij}=a_{ji}=\rho a^2$，$i\neq j=1,\cdots,k$ で ρ は既知とする．

v/σ^2 は $(u_1,\cdots,u_k)$ と独立な確率変数で，自由度 m のカイ2乗分布を持つとする．$\mathscr{C}_0$ を $\sum_i c_i=0$ かつ $c_1,\cdots,c_k$ のすべては0ではない実ベクトル $(c_1,\cdots,c_k)$ の集合とすれば，不等式

(10.5.16)
$$\sum_i c_iu_i-D\leqslant\sum_i c_i\mu_i\leqslant\sum_i c_iu_i+D$$

が $\mathscr{C}_0$ のすべてのベクトルに対して同時に成り立つ確率は γ である．ここで

(10.5.17)
$$D=R_{k,m,\gamma}\sqrt{a^2v(1-\rho)/m}\left(\frac{1}{2}\sum_i|c_i|\right).$$

$c_1,\cdots,c_k$ がすべて0の場合には（これは $\mathscr{C}_0$ に含まれないが），(10.5.16) は確率1

で成り立つ．

10.5.3 の証明には，まず $u_i-\mu_i=y_i$，$i=1,\cdots,k$ とする．このとき $(y_1,\cdots,y_k)$ は分布 $N(\{0\},\|\sigma^2a_{ij}\|)$ を持つ．ただし a_{ij} は **10.5.3** で定義されている．$w=y_1+\cdots+y_k$ および $z_i=y_i+hw$ とする．ただし h は次の2次方程式を満足する定数である（ここで，$\|a_{ij}\|$ は正定値であるから，$-1/(k-1)<\rho<1$）．

(10.5.18)
$$h^2k[1+(k-1)\rho]+2h[1+(k-1)\rho]+\rho=0.$$

よって $z_1,\cdots,z_k$ は独立で，おのおの分布 $N(0,\sigma^2a^2(1-\rho))$ を持つ．ここで次の不等式を考えよう．

(10.5.19)
$$|z_i-z_j|\leqslant H,\qquad i,j=1,\cdots,k.$$

この不等式がすべて満足されるのは $(z_1,\cdots,z_k)$ の範囲 R が次式を満たす場合であって，かつそのときに限る．

(10.5.20)
$$R\leqslant H.$$

v/σ^2 は自由度 m のカイ2乗分布を持ち，$(z_1,\cdots,z_k)$ とは独立である．さらに $(z_1,\cdots,z_k)$ が $N(0,\sigma^2a^2(1-\rho))$ からの大きさ k の標本であることを用いれば，確率変数 $R\sqrt{m/[a^2v(1-\rho)]}$ はスチューデント化された範囲 $R_{k,m}$ であることがわかる．さて $(c_1,\cdots,c_k)$ を成分のすべては0ではないが $\sum_i c_i=0$ を満たすベクトルとしよう．ここで $\sum'_i$ を $c_i>0$ となる i のすべての値にわたる和を表わすものとし，同様に $\sum''_j$ は $(-c_j)>0$ となるすべての j に関する和を表わすものとすれば，

(10.5.21)
$$\sum\nolimits'_i c_i=\sum\nolimits''_j(-c_j)=\frac{1}{2}\sum_i|c_i|=K$$

と書ける．さらに

$$\begin{aligned}\sum_i c_iz_i&=\sum\nolimits'_i c_iz_i-\sum\nolimits''_j(-c_j)z_j\\&=\frac{1}{K}\left[\sum\nolimits'_i\sum\nolimits''_j c_i(-c_j)z_i-\sum\nolimits'_i\sum\nolimits''_j c_i(-c_j)z_j\right]\\&=\frac{1}{K}\sum\nolimits'_i\sum\nolimits''_j c_i(-c_j)(z_i-z_j).\end{aligned}$$

$|z_i-z_j|\leqslant H$，$i,j=1,\cdots,k$ ならば

$$\begin{aligned}\left|\sum_i c_iz_i\right|&=\frac{1}{K}\left|\sum\nolimits'_i\sum\nolimits''_j c_i(-c_j)(z_i-z_j)\right|\\&\leqslant\frac{1}{K}\sum\nolimits'_i\sum\nolimits''_j c_i(-c_j)|z_i-z_j|\leqslant H\cdot K\end{aligned}$$

が得られる．すなわち

(10.5.22) $$\left|\sum_i c_i z_i\right| \leqslant H\cdot\left(\frac{1}{2}\sum_i |c_i|\right).$$

これは $\mathscr{C}_0$ におけるすべての実ベクトル $(c_1, \cdots, c_k)$ に対して成り立つ．しかし，(10.5.22) が $\mathscr{C}_0$ におけるすべてのベクトルに対して成り立つための必要十分条件は (10.5.19) が成り立つことである．また (10.5.19) は (10.5.20) が成り立つときかつそのときにかぎり成り立つ．したがって

(10.5.23) $$P\left(\left|\sum_i c_i z_i\right| \leqslant H\left(\frac{1}{2}\sum_i |c_i|\right);\ \mathscr{C}_0 \text{ のすべての } (c_1, \cdots, c_k) \text{ に対して}\right)$$
$$= P(R \leqslant H).$$

$R_{k,m}$ を (10.5.15) で定義されたスチューデント化された範囲とすれば

(10.5.24) $$R_{k,m} = \sqrt{\frac{R^2 m}{a^2 v(1-\rho)}}.$$

よって

(10.5.25) $$H = R_{k,m,\gamma}\sqrt{\frac{a^2 v(1-\rho)}{m}}$$

とすれば，次が得られる（ただし $R_{k,m,\gamma}$ は (10.5.15) に定義してある).

(10.5.26) $$P\left(R \leqslant R_{k,m,\gamma}\sqrt{\frac{a^2 v(1-\rho)}{m}}\right) = \gamma.$$

したがって

(10.5.27) $$P\left(\left|\sum_i c_i z_i\right| \leqslant D;\ \mathscr{C}_0 \text{ のすべての } (c_1, \cdots, c_k) \text{ に対して}\right) = \gamma.$$

ただし D は (10.5.17) で与えられている．しかし

$$\sum_i c_i z_i = \sum_i c_i(y_i + hw) = \sum_i c_i(u_i - \mu_i)$$

であるから

(10.5.28) $$P\left(\left|\sum_i c_i u_i - \sum_i c_i \mu_i\right| \leqslant D;\ \mathscr{C}_0 \text{ のすべての } (c_1, \cdots, c_k) \text{ に対して}\right)$$
$$= \gamma.$$

これは不等式 (10.5.16) が $\mathscr{C}_0$ におけるすべての $(c_1, \cdots, c_k)$ に対して同時に成り立つ確率は γ であるということと同値である．これで **10.5.3** の結論が得られた．

$\mathscr{C}_0$ のベクトルの任意有限個の組に対して (10.5.16) に対応する不等式が同時に成り立つ確率は γ を越える．たとえば $i > j = 1, \cdots, k$ のようなすべての i, j の選び方に対して不等式

$$(u_i - u_j) - R_{k,m,\gamma}\sqrt{\frac{a^2 v(1-\rho)}{m}} \leqslant (\mu_i - \mu_j)$$
$$\leqslant (u_i - u_j) + R_{k,m,\gamma}\sqrt{\frac{a^2 v(1-\rho)}{m}}$$

が同時に成り立つ確率は γ を越える．もちろんこのことはすべての $i > j = 1, \cdots, k$ に対して，

$$\left((u_i - u_j) \pm R_{k,m,\gamma}\sqrt{\frac{a^2 v(1-\rho)}{m}}\right)$$

が各 $(\mu_i - \mu_j)$ に対する信頼区間であって，すべての差 $(\mu_i - \mu_j)$, $i > j = 1, \cdots, k$ がそれぞれの各信頼区間に同時に含まれる確率は γ を越えるということと同値である．

Dwass (1959) は特別な場合としてシェフィーとトゥキーの結果を含むような同時区間に関する問題の一般化を定式化している．また，回帰直線に対する信頼領域の特別な場合における同時信頼区間の基礎的な概念が Working と Hotelling (1929) によって述べられている．

10.6 実験計画における正規線形回帰分析

本節では，10.3 節の回帰論を，モデル I 表現として知られている実験計画のごく簡単な確率的表現に適用してみよう．ここでは 2 因子，3 因子，ラテン方格の 3 計画のみを考察する．これらの計画およびその他の計画ならびに実験計画の統計的分析に深い関心を持つ読者は， Fisher (1935 *a*), Cochran と Cox (1957), Graybill (1961), Kempthorne (1946), Mann (1949), Scheffé (1959)を参照されたい．実験計画表は Fisher と Yates (1938)，北川と三留 (1953) にある．実験計画に関する文献の案内は， Greenwood と Hartley (1961) にまとめられている．

回帰分析論は現在ではいわゆる**応答面分析**にも適用されている．この型の分析に主に貢献している文献には Box (1952, 1954), Box と Hunter (1957), Box と Wilson (1951) がある．

（a）**完備2因子実験計画**

この型の実験では，rs 個の独立な確率変数の組 $\{x_{\xi\eta};\xi=1,\cdots,r,\ \eta=1,\cdots,s\}$ がすべて等分散 σ^2 を持ち，平均値

$$\mathscr{E}(x_{\xi\eta})=\mu+\mu_{\xi\cdot}+\mu_{\cdot\eta} \tag{10.6.1}$$

を持つと仮定する．ただし

$$\sum_{\xi=1}^{r}\mu_{\xi\cdot}=\sum_{\eta=1}^{s}\mu_{\cdot\eta}=0 \tag{10.6.2}$$

(10.6.1) より次のことがわかる．すべての ξ, η について，条件 (10.6.2) を満たす $\mathscr{E}(x_{\xi\eta})$ は固定変数が 0, 1, -1 のいずれかの値を持つような，(10.3.1) 型の $r+s+1$ 個の回帰係数 μ, $\mu_{\xi\cdot}$, $\mu_{\cdot\eta}$ の1次関数である．

この枠組では，r 行，s 列からなる矩形行列が考えられる．r 個の行は要因 R の水準またはカテゴリー $R_1,\cdots,R_r$ であり，s 個の行列は要因 C の所与の s 個の水準またはカテゴリー $C_1,\cdots,C_s$ である．すると $x_{\xi\eta}$ は R と C の要因の結合 (R_ξ,C_η) に対応する応答または結果を表わす確率変数である．集合 $\{x_{\xi\eta}\}$ は**応答確率変数**と呼ばれる．

注意　これらの概念をもっと具体的に説明しよう．いま s 個の機械 $C_1,\cdots,C_s$ と r 人の運転者 $R_1,\cdots,R_r$, $s\geqslant r$ がいて，ある種の製品を生産するものとしよう．各運転者が各機械を1日に8時間運転するものとし，$x_{\xi\eta}$ を運転者 R_ξ が機械 C_η を運転して産出した量とする．1日8時間運転を各機械で合計 s 回繰り返せば，2因子実験が得られる．

(10.6.1) における回帰係数 μ を全体にわたる**平均応答水準**と呼び，$\mu_{\xi\cdot}$ を R_ξ による**差効果**，$\mu_{\cdot\eta}$ を C_η による**差効果**と呼ぶ．ここで，μ, $\mu_{\xi\cdot}$, $\mu_{\cdot\eta}$ に対する最小分散線形推定量と，これらの推定量の分散および σ^2 に対する推定量を求めよう．

確率変数の組 $\{x_{\xi\eta}\}$ から，平均 $\bar{x}_{\cdot\cdot}$, $\bar{x}_{\xi\cdot}$, $\bar{x}_{\cdot\eta}$ を (8.6.7) とまったく同様に定義する．さらに

$$\begin{aligned} &m=\bar{x}_{\cdot\cdot}\\ &m_{\xi\cdot}=\bar{x}_{\xi\cdot}-\bar{x}_{\cdot\cdot}, &&\xi=1,\cdots,r\\ &m_{\cdot\eta}=\bar{x}_{\cdot\eta}-\bar{x}_{\cdot\cdot}, &&\eta=1,\cdots,s \end{aligned} \tag{10.6.3}$$

かつ

$$S=\sum_{\xi,\eta}(x_{\xi\eta}-\mu-\mu_{\xi\cdot}-\mu_{\cdot\eta})^2$$

$$S_{\cdot\cdot}=\sum_{\xi,\eta}(x_{\xi\eta}-m-m_{\xi\cdot}-m_{\cdot\eta})^2$$

$$\begin{aligned} S_{\cdot0}(\mu_{\xi\cdot})&=\sum_{\xi,\eta}(m_{\xi\cdot}-\mu_{\xi\cdot})^2\\ S_{0\cdot}(\mu_{\cdot\eta})&=\sum_{\xi,\eta}(m_{\cdot\eta}-\mu_{\cdot\eta})^2\\ S_{00}(\mu)&=\sum_{\xi,\eta}(m-\mu)^2 \end{aligned} \tag{10.6.4}$$

とする．

次式は初等的な演算で証明される．

$$S=S_{\cdot\cdot}+S_{\cdot0}(\mu_{\xi\cdot})+S_{0\cdot}(\mu_{\cdot\eta})+S_{00}(\mu) \tag{10.6.5}$$

$S_{\cdot\cdot}$ は (8.6.9) の $S_{\cdot\cdot}$ と同一であり，$S_{\cdot0}(0)=S_{\cdot0}$, $S_{0\cdot}(0)=S_{0\cdot}$ である．また S_T を (8.6.9) で与えると，$S_{\cdot\cdot}+S_{\cdot0}(0)+S_{0\cdot}(0)=S_T$ である．

10.3.1 より，μ, $\mu_{\xi\cdot}$, $\mu_{\cdot\eta}$ に対する最小分散線形推定量は (10.6.4) の S を最小にする値である．この最小化の方法によれば，(10.6.4) を吟味し，**10.3.1** および **10.3.2** を用いればわかるように，μ, $\mu_{\xi\cdot}$, $\mu_{\cdot\eta}$ に対する最小分散線形推定量として，m, $m_{\xi\cdot}$, $m_{\cdot\eta}$ が出る．さらに **10.3.3** から，$S_{\cdot\cdot}/(r-1)(s-1)$ は σ^2 の不偏推定量であることがわかる．

推定量 m, $m_{\xi\cdot}$, $m_{\cdot\eta}$ の共分散行列は次式で与えられる要素を持つ．

$$\begin{aligned} &\sigma^2(m)=\frac{\sigma^2}{rs}\\ &\sigma(m,m_{\xi\cdot})=\sigma(m,m_{\cdot\eta})=0\\ &\sigma^2(m_{\xi\cdot})=\frac{r-1}{rs}\sigma^2\\ &\sigma^2(m_{\cdot\eta})=\frac{s-1}{rs}\sigma^2\\ &\sigma(m_{\xi\cdot},m_{\xi'\cdot})=\sigma(m_{\cdot\eta},m_{\cdot\eta'})=-\frac{\sigma^2}{rs}, \qquad \xi\neq\xi',\eta\neq\eta'\\ &\sigma(m_{\xi\cdot},m_{\cdot\eta})=0 \end{aligned} \tag{10.6.6}$$

この証明は読者に残しておこう．これらの行列要素の不偏推定量は，おのおのの量において，σ^2 を $S_{\cdot\cdot}/(r-1)(s-1)$ で置き換えることによって得られる．特に，3つの組 m, $\{m_{\alpha\cdot}\}$, $\{m_{\cdot\beta}\}$ の任意の2つの m の間の共分散は0である．

結論をまとめると次のようになる．

10.6.1　$\{x_{\xi\eta};\xi=1,\cdots,r,\ \eta=1,\cdots,s\}$ を独立な確率変数の組とし，すべて等分散

σ^2 と平均 $\mathscr{E}(x_{\xi\eta})=\mu+\mu_{\xi\cdot}+\mu_{\cdot\eta}$ を持つとする．ただし $\sum_{\xi}\mu_{\xi\cdot}=\sum_{\eta}\mu_{\cdot\eta}=0$. μ, $\mu_{\xi\cdot}$, $\mu_{\cdot\eta}$ の最小分散線形推定量はそれぞれ m, $m_{\xi\cdot}$, $m_{\cdot\eta}$ であり，$S_{\cdot\cdot}$ を (10.6.4) で定義すれば，$S_{\cdot\cdot}/(r-1)(s-1)$ は σ^2 の不偏推定量である．確率変数の3つの組 m, $\{m_{\xi\cdot}\}$, $\{m_{\cdot\eta}\}$ の任意の2つの m の間の共分散は0である．すべての m の間の共分散行列は (10.6.6) で与えられる要素を持つ．

さらに，確率変数の組 $\{x_{\xi\eta}\,;\,\xi=1,\cdots,r,\ \eta=1,\cdots,s\}$ が正規分布 $N(\mu+\mu_{\xi\cdot}+\mu_{\cdot\eta},\sigma^2)$ に従って独立に分布している場合には，**10.3.4** と同様の議論により次の結果が得られる．

10.6.2　$\{x_{\xi\eta}\,;\,\xi=1,\cdots,r,\ \eta=1,\cdots,s\}$ を $\sum_{\xi}\mu_{\xi\cdot}=\sum_{\eta}\mu_{\cdot\eta}=0$ となるような分布 $N(\mu+\mu_{\xi\cdot}+\mu_{\cdot\eta},\sigma^2)$ を持つ独立な確率変数の組とする．このとき，m, $\{m_{\xi\cdot}\}$, $\{m_{\cdot\eta}\}$ は正規分布する独立な確率変数の3つの組で，それぞれ平均 μ, $\{\mu_{\xi\cdot}\}$, $\{\mu_{\cdot\eta}\}$ を持つ．また (10.6.6) で与えられる共分散行列を持ち，後の2つの分布は退化し，それぞれ制約 $\sum_{\xi}m_{\xi\cdot}=0$, $\sum_{\eta}m_{\cdot\eta}=0$ を受ける．さらに，$S_{\cdot\cdot}/\sigma^2$, $S_{\cdot 0}(\mu_{\xi\cdot})/\sigma^2$, $S_{0\cdot}(\mu_{\cdot\eta})/\sigma^2$, $S_{00}(\mu)/\sigma^2$ は独立な確率変数で，それぞれカイ2乗分布 $C((r-1)(s-1))$, $C(r-1)$, $C(s-1)$, $C(1)$ を持つ．

この結果とスチューデント分布を用いれば，定数 $\mu\,;\,\mu_{1\cdot},\cdots,\mu_{r\cdot}\,;\,\mu_{\cdot 1},\cdots,\mu_{\cdot s}$ のいずれに対しても信頼区間を求めることができる．たとえば，μ, $\mu_{\xi\cdot}$, $\mu_{\cdot\eta}$ の $100\gamma\%$ 信頼区間はそれぞれ

(10.6.7)
$$m\pm t_{(r-1)(s-1),\gamma}\cdot\sqrt{\frac{S_{\cdot\cdot}}{rs(r-1)(s-1)}}$$
$$m_{\xi\cdot}\pm t_{(r-1)(s-1),\gamma}\cdot\sqrt{\frac{S_{\cdot\cdot}}{rs(s-1)}}$$
$$m_{\cdot\eta}\pm t_{(r-1)(s-1),\gamma}\cdot\sqrt{\frac{S_{\cdot\cdot}}{rs(r-1)}}.$$

差 $\mu_{\xi\cdot}-\mu_{\xi'\cdot}$, $\xi\neq\xi'$ および μ のその他の1次関数の信頼区間も同様に定められる．また **10.5.2**, **10.5.3** を用いれば，$\mu_{\xi\cdot}-\mu_{\xi'\cdot}$, $\xi>\xi'=1,\cdots,r$ (または $\mu_{\cdot\eta}-\mu_{\cdot\eta'}$, $\eta>\eta'=1,\cdots,s$) が少なくとも確率 γ で同時に成り立つような信頼区間の組を定めることができる．

2つまたはそれ以上の μ に対する同時推定量についても (10.4.6) と同様な信頼領域を定めることができる．特に重要な場合は $\mu_{\xi\cdot}$ (または $\mu_{\cdot\eta}$) を同時に推定しようとするときである．ここでは，$\mu_{\xi\cdot}$ を同時に推定してみよう．**10.6.2** で述べたように $S_{\cdot 0}(\mu_{\xi\cdot})/\sigma^2$

と $S_{\cdot\cdot}/\sigma^2$ はカイ2乗分布 $C(r-1)$, $C((r-1)(s-1))$ に従って独立に分布するので，**7.8.5** から $(s-1)S_{\cdot 0}(\mu_{\xi\cdot})/S_{\cdot\cdot}$ はスネディッカー分布 $S((r-1),(r-1)(s-1))$ を持つことがわかる．したがって，(10.4.6) の表現を用いれば

(10.6.8)
$$P\left(\frac{(s-1)S_{\cdot 0}(\mu_{\xi\cdot})}{S_{\cdot\cdot}}<\mathscr{F}_{r-1,(r-1)(s-1),\gamma}\right)=\gamma$$

が得られる．これは

(10.6.9)
$$P((\mu_{1\cdot},\cdots,\mu_{r\cdot})\in R_\gamma)=\gamma$$

とも表わせる．ただし R_γ は次式で表わされる点 $(\mu_{1\cdot},\cdots,\mu_{r\cdot})$ に対する $(r-1)$ 次元 $100\gamma\%$ 信頼領域である．すなわち，条件 $\sum_{\xi,\eta}(\mu_{\xi\cdot}-m_{\xi\cdot})=0$ を満たす

(10.6.10)
$$\sum_{\xi,\eta}(\mu_{\xi\cdot}-m_{\xi\cdot})^2=\frac{S_{\cdot\cdot}}{s-1}\mathscr{F}_{r-1,(r-1)(s-1),\gamma}$$

のような球である．特に $\mu_{\xi\cdot}$ がすべて0になるような可能性に注目すれば，等式 (10.6.10) を持つ球が点 $(0,\cdots,0)$ を含むかどうかがわかる．これは，単に不等式

(10.6.11)
$$\frac{(s-1)S_{\cdot 0}(0)}{S_{\cdot\cdot}}<\mathscr{F}_{r-1,(r-1)(s-1),\gamma}$$

の成立を調べれば良い．これが成り立つとき，確率変数の組 $\{x_{\xi\eta}\}$ は，$\mu_{\xi\cdot}$ が $100\gamma\%$ 信頼水準ですべて0という統計的仮説を支持するという．(10.6.11) が成り立つ場合，$m_{\xi\cdot}$ ($\mu_{\xi\cdot}$ に対する推定量) は $100(1-\gamma)\%$ 有意水準で0と有意差はないといい換えてもよい．重要な点は検定に用いられる確率変数，すなわち $(s-1)S_{\cdot 0}(0)/S_{\cdot\cdot}$ が観測可能であるということである．

同様な方法で，確率変数の組 $\{x_{\xi\eta}\}$ が，$\mu_{\cdot\eta}$ がすべて0という統計的仮説を支持するか

表 10.1　完備2因子実験計画のモデルI分散分析表

要因	自由度 (D.F.)	2乗和 (S.S.)	2乗和平均 (M.S.S.)	スネディッカー $\mathscr{F}$ 比
行	$r-1$	$S_{\cdot 0}$	$\dfrac{S_{\cdot 0}}{(r-1)}$	$(s-1)S_{\cdot 0}/S_{\cdot\cdot}=\mathscr{F}_{\cdot 0}$
列	$s-1$	$S_{0\cdot}$	$\dfrac{S_{0\cdot}}{(s-1)}$	$(r-1)S_{0\cdot}/S_{\cdot\cdot}=\mathscr{F}_{0\cdot}$
残差 (誤差)	$(r-1)(s-1)$	$S_{\cdot\cdot}$	$\dfrac{S_{\cdot\cdot}}{(r-1)(s-1)}$	
計	$rs-1$	S_T		

どうか調べられる．

$S_{\cdot 0}(0) = S_{\cdot 0}$，$S_{0\cdot}(0) = S_{0\cdot}$ を考慮して，正規性という仮定の下で完備2因子実験計画のモデルI分散分析表を，表10.1のように表わそう．

1番上のスネディッカー比 $\mathscr{F}_{\cdot 0}$ は $\mu_{\xi\cdot}$ がすべて0という仮説を検定するためであり，2番目の比は，$\mu_{\cdot\eta}$ がすべて0という仮説を検定するものである．分散分析の構成を，表10.1のような形にまとめる方法は Fisher (1925 *a*) による．また彼はモデルI分散分析法の理論およびその応用も発展させた．

(b)　完備3因子実験計画

10.6 (a) 節の概念はそのまま高次元計画，すなわち，3またはそれ以上の因子（要因）を含む実験計画に拡張することができる．3因子への拡張を調べてみよう．ここで，rst 個の独立な確率変数の組 $\{x_{\xi\eta\zeta};\ \xi = 1, \cdots, r,\ \eta = 1, \cdots, s,\ \zeta = 1, \cdots, t\}$ の分散はすべて σ^2 に等しく，平均値は

(10.6.12)　$$\mathscr{E}(x_{\xi\eta\zeta}) = \mu + \mu_{\xi\cdot\cdot} + \mu_{\cdot\eta\cdot} + \mu_{\cdot\cdot\zeta} + \mu_{\xi\eta\cdot} + \mu_{\xi\cdot\zeta} + \mu_{\cdot\eta\zeta}$$

で与えられるとする．ただし

$$\sum_{\xi} \mu_{\xi\cdot\cdot} = \sum_{\eta} \mu_{\cdot\eta\cdot} = \sum_{\zeta} \mu_{\cdot\cdot\zeta} = 0, \qquad \sum_{\xi} \mu_{\xi\eta\cdot} = \sum_{\eta} \mu_{\xi\eta\cdot} = 0,$$

$$\sum_{\xi} \mu_{\xi\cdot\zeta} = \sum_{\zeta} \mu_{\xi\cdot\zeta} = 0, \qquad \sum_{\eta} \mu_{\cdot\eta\zeta} = \sum_{\zeta} \mu_{\cdot\eta\zeta} = 0$$

この型の実験計画の配置では，r 個の行，s 個の列，t 個の層を考え，行は要因 R の水準 $R_1, \cdots, R_r$，列は要因 C の水準 $C_1, \cdots, C_s$，層は要因 L の水準 $L_1, \cdots, L_t$ と考えられる．定数 $\mu_{\xi\cdot\cdot}$ は R_ξ による差効果であり，$\mu_{\cdot\eta\cdot}$，$\mu_{\cdot\cdot\zeta}$ についても同様な解釈ができる．一方 $\mu_{\xi\eta\cdot}$ は (R_ξ, C_η) あるいは R_ξ と C_η の間の交互作用による差効果であり，$\mu_{\xi\cdot\zeta}$ および $\mu_{\cdot\eta\zeta}$ についても同様である．

種々の μ を推定するために，(8.6.7) を拡張して，$\bar{x}_{\cdots}$，$\bar{x}_{\xi\cdot\cdot}$，$\bar{x}_{\cdot\eta\cdot}$，$\bar{x}_{\cdot\cdot\zeta}$，$\bar{x}_{\xi\eta\cdot}$，$\bar{x}_{\xi\cdot\zeta}$，$\bar{x}_{\cdot\eta\zeta}$ を定義する．また

$$m = \bar{x}_{\cdots}$$

(10.6.13)　$$m_{\xi\cdot\cdot} = \bar{x}_{\xi\cdot\cdot} + \bar{x}_{\cdots}, \qquad m_{\cdot\eta\cdot}, m_{\cdot\cdot\zeta}\ \text{についても同様}$$

$$m_{\xi\eta\cdot} = \bar{x}_{\xi\eta\cdot} - \bar{x}_{\xi\cdot\cdot} - \bar{x}_{\cdot\eta\cdot} + \bar{x}_{\cdots}, \qquad m_{\xi\cdot\zeta}, m_{\cdot\eta\zeta}\ \text{についても同様}$$

を定義する．さらに，

(10.6.14)

$$S = \sum_{\xi,\eta,\zeta} (x_{\xi\eta\zeta} - \mu - \mu_{\xi\cdot\cdot} - \mu_{\cdot\eta\cdot} - \mu_{\cdot\cdot\zeta} - \mu_{\xi\eta\cdot} - \mu_{\xi\cdot\zeta} - \mu_{\cdot\eta\zeta})^2$$

$$S_{\cdots} = \sum_{\xi,\eta,\zeta} (x_{\xi\eta\zeta} - m - m_{\xi\cdot\cdot} - m_{\cdot\eta\cdot} - m_{\cdot\cdot\zeta} - m_{\xi\eta\cdot} - m_{\xi\cdot\zeta} - m_{\cdot\eta\zeta})^2$$

$$S_{\cdot\cdot 0}(\mu_{\xi\eta\cdot}) = \sum_{\xi,\eta,\zeta} (m_{\xi\eta\cdot} - \mu_{\xi\eta\cdot})^2, \qquad S_{\cdot 0\cdot}(\mu_{\xi\cdot\zeta})\ \text{および}\ S_{0\cdot\cdot}(\mu_{\cdot\eta\zeta})\ \text{についても同様}$$

$$S_{\cdot 00}(\mu_{\xi\cdot\cdot}) = \sum_{\xi,\eta,\zeta} (m_{\xi\cdot\cdot} - \mu_{\xi\cdot\cdot})^2, \qquad S_{0\cdot 0}(\mu_{\cdot\eta\cdot})\ \text{および}\ S_{00\cdot}(\mu_{\cdot\cdot\zeta})\ \text{についても同様}$$

$$S_{000}(\mu) = \sum_{\xi,\eta,\zeta} (m - \mu)^2$$

とする．$S_{\cdots}$ は (8.6.31) の定義とまったく同じであるが，$S_{\cdot\cdot 0}(0) = S_{\cdot\cdot 0}$，$S_{\cdot 00}(0) = S_{\cdot 00}$ である．ただし，$S_{\cdot\cdot 0}$，$S_{\cdot 00}$ は (8.6.31) で定義されている．このとき S は

(10.6.15)　$$S = S_{\cdots} + S_{\cdot\cdot 0}(\mu_{\xi\eta\cdot}) + S_{\cdot 0\cdot}(\mu_{\xi\cdot\zeta}) + S_{0\cdot\cdot}(\mu_{\cdot\eta\zeta}) + S_{\cdot 00}(\mu_{\xi\cdot\cdot}) + S_{0\cdot 0}(\mu_{\cdot\eta\cdot}) + S_{00\cdot}(\mu_{\cdot\cdot\zeta}) + S_{000}(\mu)$$

と分割できる．**10.3.1** を用いれば，μ，$\mu_{\xi\cdot\cdot}$，$\mu_{\cdot\eta\cdot}$，$\mu_{\cdot\cdot\zeta}$，$\mu_{\xi\eta\cdot}$，$\mu_{\xi\cdot\zeta}$，$\mu_{\cdot\eta\zeta}$ の最小分散線形推定量はそれぞれ，m，$m_{\xi\cdot\cdot}$，$m_{\cdot\eta\cdot}$，$m_{\cdot\cdot\zeta}$，$m_{\xi\eta\cdot}$，$m_{\xi\cdot\zeta}$，$m_{\cdot\eta\zeta}$ である．さらに **10.3.3** より，$S_{\cdots}/[(r-1)(s-1)(t-1)]$ が σ^2 の不偏推定量となる．

m の共分散行列の中には次の非0要素だけを記入すれば十分である．残りの非0要素は対称性を考慮すればすぐに求められる．

$$\sigma^2(m) = \frac{\sigma^2}{rst}, \qquad \sigma^2(m_{\xi\cdot\cdot}) = \frac{(r-1)\sigma^2}{rst}$$

$$\sigma(m_{\xi\cdot\cdot}, m_{\xi'\cdot\cdot}) = -\frac{\sigma^2}{rst}, \qquad \xi \neq \xi'$$

(10.6.16)　$$\sigma^2(m_{\xi\eta\cdot}) = \frac{(r-1)(s-1)\sigma^2}{rst},$$

$$\sigma(m_{\xi\eta\cdot}, m_{\xi\eta'\cdot}) = \frac{-(r-1)\sigma^2}{rst}, \qquad \eta \neq \eta'$$

$$\sigma(m_{\xi\eta\cdot}, m_{\xi'\eta'\cdot}) = -\frac{\sigma^2}{rst}, \qquad \xi \neq \xi', \eta \neq \eta'$$

組 m，$\{m_{\xi\cdot\cdot}\}$，$\{m_{\cdot\eta\cdot}\}$，$\{m_{\cdot\cdot\zeta}\}$，$\{m_{\xi\eta\cdot}\}$，$\{m_{\xi\cdot\zeta}\}$，$\{m_{\cdot\eta\zeta}\}$ の中の任意の2つの m の間の共分散は0である．

(10.6.16) の σ^2 のかわりに $S_{\cdots}/[(r-1)(s-1)(t-1)]$ を代入すれば，(10.6.16) で定義された分散および共分散に対する不偏推定量が得られる．

3因子実験計画に関する **10.6.1**, **10.6.2** と同様な式は読者自身求めよ．

10.6.2 を3因子実験計画に拡張する場合に注意することは，組 $\{x_{\xi\eta\zeta}\}$ の確率変数が独立で，正規分布

(10.6.17) $$N(\mu+\mu_{\xi\cdot\cdot}+\mu_{\cdot\eta\cdot}+\mu_{\cdot\cdot\zeta}+\mu_{\xi\eta\cdot}+\mu_{\xi\cdot\zeta}+\mu_{\cdot\eta\zeta},\sigma^2)$$

を持てば

(10.6.18) $$\frac{S_{\cdots}}{\sigma^2},\quad \frac{S_{\cdot\cdot0}(\mu_{\xi\eta\cdot})}{\sigma^2},\quad \frac{S_{\cdot0\cdot}(\mu_{\xi\cdot\zeta})}{\sigma^2},\quad \frac{S_{0\cdot\cdot}(\mu_{\cdot\eta\zeta})}{\sigma^2}$$
$$\frac{S_{\cdot00}(\mu_{\xi\cdot\cdot})}{\sigma^2},\quad \frac{S_{0\cdot0}(\mu_{\cdot\eta\cdot})}{\sigma^2},\quad \frac{S_{00\cdot}(\mu_{\cdot\cdot\zeta})}{\sigma^2},\quad \frac{S_{000}(\mu)}{\sigma^2}$$

はそれぞれ自由度が

(10.6.19) $$(r-1)(s-1)(t-1),\quad (r-1)(s-1),\quad (r-1)(t-1),$$
$$(s-1)(t-1),\quad (r-1),\quad (s-1),\quad (t-1),\quad 1$$

のカイ2乗分布に従って独立に分布することである．ここで，2因子実験計画に関する (10.6.7) と同様に個々の μ に対する信頼区間を，あるいは (10.6.10) と同様に，いくつかの μ の同時推定量に対する信頼領域を定めることができる．**10.5.2** または **10.5.3** を用いれば，$\{\mu_{\xi\cdot\cdot}-\mu_{\xi'\cdot\cdot}\,;\,\xi>\xi'=1,\cdots,r\}$，$\{\mu_{\xi\eta\cdot}-\mu_{\xi'\eta\cdot}\,;\,\xi>\xi'=1,\cdots,r$ で η は固定$\}$ のような差の組に対する同時信頼区間も求めることができる．完備3因子実験計画に関する表10.1と同じモデルI分散分析表は読者自身で求めてみよ．

(c) ラテン方格実験計画法

上で述べた完備3因子実験計画法は応用次第では，種々の実験計画を構成できる確率変数を含んでいる．多くの実験計画は釣合い不完備3因子実験計画法に帰着できる．この種の最も簡単な実験計画は，ラテン方格実験計画である．ここではこの計画を考察しよう．これは，r 行 r 列 r 層からなる3次元標本行列 $\{x_{\xi\eta\zeta}\}$ から応答確率変数の釣合い不完備行列標本 $\{x_{\xi\eta\zeta}\}^*$ を選ぶものである．

ラテン方格計画では，3因子 R, C, L を考え，各因子に r 個の水準，$R_1,\cdots,R_r$；$C_1,\cdots,C_r$；$L_1,\cdots,L_r$ を設ける．3次元の矩形配置をした胞体において R_ξ は ξ 番目の行を，C_η は η 番目の列を，L_ζ は ζ 番目の層を表わす．ここで考えるラテン方格の確率変数の標本 $\{x_{\xi\eta\zeta}\}^*$ は，この $r\times r\times r$ 個配列からなる行列標本 $\{x_{\xi\eta\zeta}\}$ からの r^2 個の成分を持つ釣合い不完備行列標本である．すなわち $\{x_{\xi\eta\zeta}\}^*$ はそれぞれ (ξ,η), (η,ζ), (ζ,ξ) なる値の各組合せにただ1つの確率変数が存在するように選ばれる［8.6(d)節を見よ］．このように選ばれた (ξ,η,ζ) の r^2 個の値の集合を G^* としよう．

さて標本 $\{x_{\xi\eta\zeta}\}^*$ に含まれる r^2 個の要素は独立で正規分布 $N(\mathscr{E}(x_{\xi\eta\zeta}),\sigma^2)$ を持つとする．ここで任意の $(\xi,\eta,\zeta)\in G^*$ に対して

(10.6.20) $$\mathscr{E}(x_{\xi\eta\zeta})=\mu+\mu_{\xi\cdot\cdot}+\mu_{\cdot\eta\cdot}+\mu_{\cdot\cdot\zeta}$$

とする．ただし

(10.6.21) $$\sum_{\xi}\mu_{\xi\cdot\cdot}=\sum_{\eta}\mu_{\cdot\eta\cdot}=\sum_{\zeta}\mu_{\cdot\cdot\zeta}=0$$

これを (10.6.12) と比較して，特に注目すべきことは，2次交互作用 $\mu_{\xi\eta\cdot}$, $\mu_{\xi\cdot\zeta}$, $\mu_{\cdot\eta\zeta}$ はすべて0と仮定していることである．

ラテン方格は (10.6.20) の各 μ に対する比較的簡単な最小分散線形推定量を与える．これを見るために，$\bar{x}^*_{\cdots}$, $\bar{x}^*_{\xi\cdot\cdot}$, $\bar{x}^*_{\cdot\eta\cdot}$, $\bar{x}^*_{\cdot\cdot\zeta}$ を (8.6.40) のように定義されたものとする．また m^*, $m^*_{\xi\cdot\cdot}$, $m^*_{\cdot\eta\cdot}$, $m^*_{\cdot\cdot\zeta}$ は，(10.6.13) で m, $m_{\xi\cdot\cdot}$, $m_{\cdot\eta\cdot}$, $m_{\cdot\cdot\zeta}$ が $\{x_{\xi\eta\zeta}\}$ に含まれる x によって定義されたのと同様に，$\{x_{\xi\eta\zeta}\}^*$ に含まれる x によって定義されるものとしよう．さらに $S^*_{\cdot00}(\mu_{\xi\cdot\cdot})$, $S^*_{0\cdot0}(\mu_{\cdot\eta\cdot})$, $S^*_{00\cdot}(\mu_{\cdot\cdot\zeta})$, $S^*_{000}(\mu)$ は(10.6.14)の m のかわりに m^* で置き換えて定義する．もちろん和の範囲は $(\xi,\eta,\zeta)\in G^*$ すべてにおよぶ．また

(10.6.22) $$S^*=\sum_{\xi,\eta,\zeta}{}^*(x_{\xi\eta\zeta}-\mu-\mu_{\xi\cdot\cdot}-\mu_{\cdot\eta\cdot}-\mu_{\cdot\cdot\zeta})^2$$
$$S^*_{\cdots}=\sum_{\xi,\eta,\zeta}{}^*(x_{\xi\eta\zeta}-m^*-m^*_{\xi\cdot\cdot}-m^*_{\cdot\eta\cdot}-m^*_{\cdot\cdot\zeta})^2$$

とする．ただし $\sum\limits_{\xi,\eta,\zeta}{}^*$ は (8.6.40) と同様に，すべての $(\xi,\eta,\zeta)\in G^*$ に関する和を表わす．

この S^* は次のように分割される．

(10.6.23) $$S^*=S^*_{\cdots}+S^*_{\cdot00}(\mu_{\xi\cdot\cdot})+S^*_{0\cdot0}(\mu_{\cdot\eta\cdot})+S^*_{00\cdot}(\mu_{\cdot\cdot\zeta})+S^*_{000}(\mu)$$

ここに，右辺の成分はそれぞれ自由度 $(r-1)(r-2)$, $(r-1)$, $(r-1)$, $(r-1)$, 1 を持つ．**10.3.2** を用いれば，m^*, $m^*_{\xi\cdot\cdot}$, $m^*_{\cdot\eta\cdot}$, $m^*_{\cdot\cdot\zeta}$ がそれぞれ μ, $\mu_{\xi\cdot\cdot}$, $\mu_{\cdot\eta\cdot}$, $\mu_{\cdot\cdot\zeta}$ の最小分散線形推定量であることは明らかである．さらに $S^*_{\cdots}/(r-1)(r-2)$ は σ^2 の不偏推定量である．

推定量 m^*, $m^*_{\xi\cdot\cdot}$, $m^*_{\cdot\eta\cdot}$, $m^*_{\cdot\cdot\zeta}$ の共分散行列は次のような非0要素を持つ．

$$\sigma^2(m^*)=\frac{\sigma^2}{r^2},$$

(10.6.24)
$$\sigma^2(m^*_{\xi\cdot\cdot}) = \sigma^2(m^*_{\cdot\eta\cdot}) = \sigma^2(m^*_{\cdot\cdot\zeta}) = \frac{r-1}{r^2}\sigma^2,$$

$$\sigma(m^*_{\xi\cdot\cdot}, m^*_{\xi'\cdot\cdot}) = \sigma(m^*_{\cdot\eta\cdot}, m^*_{\cdot\eta'\cdot}) = \sigma(m^*_{\cdot\cdot\zeta}, m^*_{\cdot\cdot\zeta'}) = -\frac{\sigma^2}{r^2},$$

$$\xi \neq \xi', \qquad \eta \neq \eta', \qquad \zeta \neq \zeta'$$

これらの分散，共分散の不偏推定量は，σ^2 を $S^*_{\cdot\cdot}/(r-1)(r-2)$ で置き換えることによって得られる．

ラテン方格に関する **10.6.1**，**10.6.2** と同様な式へのまとめは読者にゆだねる．またラテン方格に関する表10.1と同じモデルI分散分析表も読者自身で求められたい．

10.7　線形結合による分散成分の推定

統計学の問題，特に実験計画，誤差論，心理学的要因分析法のような分野においては，0の共分散を持つある確率変数の線形結合は観測可能であるが，確率変数自身は観測不可能な場合がある．このような確率変数の線形結合の例として，(8.6.3) が掲げられる．そこでは，$x(u, v)$ は観測可能な確率変数であり，μ と観測不可能な確率変数 $\varepsilon_{u\cdot}$，$\varepsilon_{\cdot v}$，ε_{uv} の線形関数として表わされている．

このような場合，基本的な問題はこれら観測不可能な確率変数からなる観測可能の線形結合から成分確率変数の分散の推定量を求めることである．

厳密に述べよう．まず，x_ξ, $\xi = 1, \cdots, n$ を次の形をした既知確率変数とする．

(10.7.1)
$$x_\xi = \mu + \sum_{g=1}^{m} a_{g\xi}\varepsilon_{g\xi}, \qquad \xi = 1, \cdots, n \geqq m$$

ここで

(10.7.2)
$$\mathscr{E}(\varepsilon_{g\xi}) = 0, \qquad \mathscr{E}(\varepsilon^2_{g\xi}) = \sigma^2_g, \qquad g = 1, \cdots, m, \ \xi = 1, \cdots, n$$
$$\mathscr{E}(\varepsilon_{g\xi}\varepsilon_{g\xi'}) = \mathscr{E}(\varepsilon_{g\xi}\varepsilon_{g'\xi}) = 0, \qquad g \neq g', \qquad \xi \neq \xi'$$

で，$a_{g\xi}$ は既知定数で $\|a_{g\xi}\|$ の階数を m とする．$\varepsilon_{g\xi}$ の観測可能性は仮定しないとする．すなわち $\varepsilon_{g\xi}$ は未知のパラメータの関数としてよい．

さて，標本平均 $\bar{x}$ が μ の不偏推定量となることは明白である．この推定量の分散は

(10.7.3)
$$\sigma^2(\bar{x}) = \frac{1}{n^2}\sum_g \sum_\xi a^2_{g\xi}\sigma^2_g$$

になる．また s^2 を標本分散とすれば

(10.7.4)
$$\mathscr{E}(s^2) = \frac{1}{n}\sum_g \sum_\xi a^2_{g\xi}\sigma^2_g.$$

s^2 が $\frac{1}{n}\sum_g \sum_\xi a^2_{g\xi}\sigma^2_g$ の不偏推定量であることに注目すれば，s^2/n は $\sigma^2(\bar{x})$ の不偏推定量である．しかし標本 $(x_1, \cdots, x_m)$ から各分散成分 $\sigma^2_1, \cdots, \sigma^2_m$ を推定したい．このために平方和

(10.7.5)
$$S_T = \sum_\xi (x_\xi - \bar{x})^2$$

を考えよう．この右辺は，確率変数 $(x_1 - \bar{x}), \cdots, (x_n - \bar{x})$ に関する2次形式であり，これは $(x_1 - \mu), \cdots, (x_n - \mu)$ に関する階数 $n-1$ の行列を持つ2次形式として表わされる．$n - 1 \geqq m$ であれば，半正定値2次形式の基礎理論を用いて，S_T が次のように（種々の方法で）分割できることが示される．

(10.7.6)
$$S_T = S_1 + \cdots + S_m.$$

ここで $S_1, \cdots, S_m$ は $(x_1 - \bar{x}), \cdots, (x_n - \bar{x})$ の半正定値2次形式であり，それぞれ階数 $n_1, \cdots, n_m$（すべて $\geqq 1$）の行列を持つような $(x_1 - \mu), \cdots, (x_n - \mu)$ の2次形式として表わすことができる．ただし $n_1 + \cdots + n_m = n - 1$ である．このとき $h = 1, \cdots, m$ に対して

(10.7.7)
$$S_h = \sum_{\xi,\eta} A^{(h)}_{\xi\eta}(x_\xi - \mu)(x_\eta - \mu)$$

と書ける．ただし $\|A^{(h)}_{\xi\eta}\|$ は階数 n_h の対称行列であり，その要素は既知である．(10.7.1) から，$(x_\xi - \mu)$，$(x_\eta - \mu)$ を求めて (10.7.7) に代入して平均値をとれば

(10.7.8)
$$\mathscr{E}(S_h) = \sum_g B_{hg}\sigma^2_g.$$

ただし

(10.7.9)
$$B_{hg} = \sum_\xi A^{(h)}_{\xi\xi}a^2_{g\xi}.$$

さて，$\|A^{(h)}_{\xi\eta}\|$ の性質と $\|a_{g\xi}\|$ の仮定からして，$\|B_{hg}\|$ は正則であることがわかる．したがって **10.1.1** を用いれば，(10.7.8) から，$\sigma^2_1, \cdots, \sigma^2_m$ に対する次のような不偏推定量が得られる．すなわち

(10.7.10)
$$\mathscr{E}^{-1}(\sigma^2_g) = \sum_h B^{gh}S_h.$$

$\sigma^2(\bar{x})$ の不偏推定量は (10.7.3) において σ^2_g を $\mathscr{E}^{-1}(\sigma^2_g)$，$g = 1, \cdots, m$ で置き換えれば得

られる.

以上をまとめると,

10.7.1　$x_\xi,\ \xi=1,\cdots,n \geqslant m$ を次のような観測可能な確率変数とする.

$$x_\xi=\mu+\sum_{g=1}^{m} a_{g\xi}\varepsilon_{g\xi},$$

$\varepsilon_{g\xi}$ は平均, 共分散がすべて0で, $\mathscr{E}(\varepsilon_{g\xi}^2) \doteqdot \sigma_g^2$ となる確率変数である. ただし, μ と σ_g^2 は未知パラメータであり, $a_{g\xi}$ は $\|a_{g\xi}\|$ の階数が m となる既知定数である. $\bar{x}$ を $(x_1,\cdots,x_n)$ の平均, $S_T, S_1,\cdots,S_m$ を (10.7.5), (10.7.6) で定義したような $(x_1-\bar{x}),\cdots,(x_n-\bar{x})$ の2次形式とする. このとき $\bar{x}$ は μ に対する不偏推定量となり,

$$\mathscr{E}^{-1}(\sigma_g^2)=\sum_h B^{gh}S_h$$

は $\sigma_g^2,\ g=1,\cdots,m$ の不偏推定量となる. ただし

$$\|B^{gh}\|=\left\|\sum_\xi A_{\xi\xi}^{(h)}a_{g\xi}^2\right\|^{-1}$$

で, $\|A_{\xi\eta}^{(h)}\|$ は2次形式 S_h の行列である.

分散成分を推定する特殊な問題では, 既知定数の行列 $\|a_{g\xi}\|$ がただちに S_T に対する分割 (10.7.6) を暗示するような特殊形をとっている. 次節でこの原理を実験計画法への応用した場合を考えてみよう.

10.8　実験計画における分散成分の推定量

本節では, 分散成分の推定に関する概念を, いくつかの簡単な実験計画に適用してみる. 分散成分分析という見地から, 完備2因子計画, 釣合い不完備2因子計画, 完備3因子計画, ラテン方格計画のみを考えることにしよう. さらに, この問題に興味を持った読者は, Cochran と Cox (1957), Graybill (1961), Kempthorne (1952), Scheffé (1959) 等の本を参照されたい. 本節では, これらの実験計画に含まれる分散成分の不偏推定量を定める問題のみを考え, 推定量の分散までは考えないことにする. この後者の問題に関心を持つ読者は, Hooke (1956 *b*), Scheffé (1959), Tukey (1956 *b*, 1957 *a*, 1957 *b*) の論文

を参照されたい. 分散成分の推定量の信頼区間に関する結果は, Bulmer (1957), Moriguti (1954), Scheffé (1959) 等によって得られている.

確率変数の観測可能な1次関数を用いた実験計画の確率的な表現を, 10.6節で述べた正規回帰分析に基づくモデルI表示と対照させて, モデルII表示と呼ぶ. この2つの分散分析法に対して, モデルIおよびモデルII表示と命名したのは Eisenhart (1947) である.

(a)　完備2因子実験計画

完備2因子実験計画のモデルIIによる確率的表現に必要とされる基本的な標本論は 8.6 (a) 節で述べたような2次行列標本の理論ある. したがってここでは, 8.6 (a) 節での記号を用いる.

まず完備2因子実験計画へのモデルIによる接近法とモデルIIによる接近法との間の根本的な違いを述べておこう. 10.4 (a) 節で述べた完備2因子実験計画のモデルI表示では, R 要因の水準 $R_1,\cdots,R_r$ と C 要因の水準 $C_1,\cdots,C_s$ は応答確率変数 $x_{\xi\eta}$ で記述されるすべての実験全体に対して固定されていた. しかし, 状況によっては, r 個の R 要因の水準, s 個の C 要因の水準またはその両方が, 有限あるいは無限であるかも知れないような, より大きな水準の母集団からとった標本である場合もある. この場合には, 回帰分析 (モデルI) による接近法は適当ではない. ここで, 水準の母集団が無限である場合を考えてみよう. 厳密に述べれば, モデルIIによる接近法においては, R 要因の任意の水準 u と C 要因の任意の水準 v に対する応答確率変数 $x(u,v)$ は (8.6.3) で与えた形をとるものとする. この分散 σ^2 は (8.6.4) で与えられる. あるいは簡単に (8.6.4 *a*), すなわち

(10.8.1)
$$\sigma^2=\sigma_{\cdot 0}^2+\sigma_{0\cdot}^2+\sigma_{\cdot\cdot}^2.$$

で表わされる. ただし, $\sigma_{\cdot 0}^2$ は行成分 (R 要因), $\sigma_{0\cdot}^2$ は列成分 (C 要因) そして $\sigma_{\cdot\cdot}^2$ は残差成分である.

r 個の R 要因の水準が無作為にとられ, s 個の C 要因の水準が無作為に取り出されるとすれば——ともに無限母集団からとるとして——応答確率変数は 8.6 (a) 節で定義した行列標本 $\{x_{\xi\eta}\,;\,\xi=1,\cdots,r,\ \eta=1,\cdots,s\}$ となる. ただし, $x_{\xi\eta}$ は独立で, (8.6.8) すなわち

(10.8.2)
$$x_{\xi\eta}=\mu+\varepsilon_{\xi\cdot}+\varepsilon_{\cdot\eta}+\varepsilon_{\xi\eta}$$

によって表現できる. ただし μ は母集団平均であり, $\varepsilon_{\xi\cdot},\ \varepsilon_{\cdot\eta},\ \varepsilon_{\xi\eta}$ は平均0, 共分散0

を持つ確率変数である.

$x_{\xi\eta}$ の分散 σ^2 は (10.8.1) を満たしている. モデル II 分析の主たる問題は標本 $\{x_{\xi\eta}\}$ から $\sigma^2_{\cdot 0}$, $\sigma^2_{0\cdot}$ および $\sigma^2_{\cdot\cdot}$ を推定することである.

このために, (8.6.7) での $\bar{x}_{\cdot\cdot}$, $\bar{x}_{\xi\cdot}$, $\bar{x}_{\cdot\eta}$ の定義と (8.6.9) での $S_{\cdot 0}$, $S_{0\cdot}$, $S_{\cdot\cdot}$ の定義を用いる. このとき **8.6.2** の仮定の下では, (8.6.11) で述べたように

(10.8.3)
$$\mathscr{E}(S_{\cdot 0}) = (r-1)(s\sigma^2_{\cdot 0} + \sigma^2_{\cdot\cdot})$$
$$\mathscr{E}(S_{0\cdot}) = (s-1)(r\sigma^2_{0\cdot} + \sigma^2_{\cdot\cdot})$$
$$\mathscr{E}(S_{\cdot\cdot}) = (r-1)(s-1)\sigma^2_{\cdot\cdot}$$

が得られる. したがって **10.1.1** (または **10.7.1**) を用いれば, $\sigma^2_{\cdot 0}$, $\sigma^2_{0\cdot}$, $\sigma^2_{\cdot\cdot}$ の不偏推定量が次のように得られる.

(10.8.4)
$$\mathscr{E}^{-1}(\sigma^2_{\cdot 0}) = \frac{1}{s(r-1)}\left[S_{\cdot 0} - \frac{S_{\cdot\cdot}}{s-1}\right]$$
$$\mathscr{E}^{-1}(\sigma^2_{0\cdot}) = \frac{1}{r(s-1)}\left[S_{0\cdot} - \frac{S_{\cdot\cdot}}{r-1}\right]$$
$$\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot}) = \frac{S_{\cdot\cdot}}{(r-1)(s-1)}$$

母集団平均 μ に対する推定量は $\bar{x}_{\cdot\cdot}$ であるが, この分散は, (8.6.12) で述べたように $\sigma^2_{\cdot 0}/r + \sigma^2_{0\cdot}/s + \sigma^2_{\cdot\cdot}/rs$ で与えられる. よってこの分散の不偏推定量は, $\sigma^2_{\cdot 0}$, $\sigma^2_{0\cdot}$, $\sigma^2_{\cdot\cdot}$ をそれぞれ (10.8.4) の推定量で置き換えることによって得られる.

モデル II 表示の構成も通常, 表 10.2 のような分散分析表で表わす.

表 10.2　完備2因子実験計画に対するモデル II 分散分析表

要因	自由度 (D.F.)	2乗和 (S.S.)	2乗和平均 (M.S.S.)	2乗和平均の期待値
行	$r-1$	$S_{\cdot 0}$	$\dfrac{S_{\cdot 0}}{(r-1)}$	$s\sigma^2_{\cdot 0} + \sigma^2_{\cdot\cdot}$
列	$s-1$	$S_{0\cdot}$	$\dfrac{S_{0\cdot}}{(s-1)}$	$r\sigma^2_{0\cdot} + \sigma^2_{\cdot\cdot}$
残差 (誤差)	$(r-1)(s-1)$	$S_{\cdot\cdot}$	$\dfrac{S_{\cdot\cdot}}{(r-1)(s-1)}$	$\sigma^2_{\cdot\cdot}$
計	$rs-1$	S_T		

最後に次のことを注意しておこう. R 要因の水準および C 要因の水準の母集団が有限である場合には, 分散成分すなわち $\sigma^2_{f\cdot 0}$, $\sigma^2_{f0\cdot}$, $\sigma^2_{f\cdot\cdot}$ の有限母集団に対する推定量は式 (10.8.3) を (8.6.22) で置き換え, **10.7.1** を用いれば得られる. もちろん, この場合にも, 表 10.2 と同様な表をつくることができる.

(b)　釣合い不完備2因子実験計画

釣合い不完備2因子実験計画に関するモデル II の処理に含まれる標本論は 8.6(d) 節で与えられている. この計画では, 8.6(d) 節のように, 行列標本 $\{x_{\xi\eta}\}$ から大きさ $n = rs' = r's$ の応答確率変数からなる部分標本 $\{x_{\xi\eta}\}^*$ を選び出す. $\{x_{\xi\eta}\}^*$ に含まれるおのおのの要素は (10.8.2) で述べたように, それ自身が確率変数の1次結合であるような確率変数である. さらに $\{x_{\xi\eta}\}^*$ のおのおのの要素 $x_{\xi\eta}$ の分散は σ^2 で, かつ σ^2 は (10.8.1) で述べたように3成分の和である.

釣合い不完備2因子実験計画のモデル II 分析の主目的は, 部分標本 $\{x_{\xi\eta}\}^*$ の要素を用いて, μ および分散成分 $\sigma^2_{\cdot 0}$, $\sigma^2_{0\cdot}$, $\sigma^2_{\cdot\cdot}$ の不偏推定量を求めることである.

標本 $\{x_{\xi\eta}\}^*$ の平均 $\bar{x}^*_{\cdot\cdot}$ は, (8.6.39) で見たように, μ の不偏推定量であり, 次の分散を持つ.

(10.8.5)
$$\sigma^2(\bar{x}^*_{\cdot\cdot}) = \frac{\sigma^2_{\cdot\cdot}}{n} + \frac{\sigma^2_{\cdot 0}}{r} + \frac{\sigma^2_{0\cdot}}{s}$$

(8.6.39) の最後の3つの等式から, それらの等式は $r' > 1$, $s' > 1$ ならば, $\sigma^2_{\cdot 0}$, $\sigma^2_{0\cdot}$, $\sigma^2_{\cdot\cdot}$ に対する一意解を持つことがわかる. この条件の下で **10.1.1** (または **10.7.1**) を適用すれば, $\sigma^2_{\cdot 0}$, $\sigma^2_{0\cdot}$, $\sigma^2_{\cdot\cdot}$ に対する次の推定量が得られる.

(10.8.6)
$$\mathscr{E}^{-1}(\sigma^2_{\cdot 0}) = \frac{S^*_{\cdot 0} + S^*_E}{n-s} - \mathscr{E}^{-1}(\sigma^2_{\cdot\cdot})$$
$$\mathscr{E}^{-1}(\sigma^2_{0\cdot}) = \frac{S^*_{0\cdot} + S^*_E}{n-r} - \mathscr{E}^{-1}(\sigma^2_{\cdot\cdot})$$
$$\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot}) = \frac{1}{K}\left[AS^*_{0\cdot} + BS^*_{\cdot 0} + CS^*_E\right]$$

ここに

$$A = \frac{(r-r')}{(n-r)}, \qquad B = \frac{(s-s')}{(n-s)}$$
$$C = A + B - 1, \qquad K = n - r - s + 1$$

$\sigma^2(\bar{x}^*_{\cdot\cdot})$ に対する不偏推定量は, (10.8.5) の $\sigma^2_{\cdot 0}$, $\sigma^2_{0\cdot}$, $\sigma^2_{\cdot\cdot}$ にそれぞれ推定量 $\mathscr{E}^{-1}(\sigma^2_{\cdot 0})$, $\mathscr{E}^{-1}(\sigma^2_{0\cdot})$, $\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot})$ を代入すれば得られる.

不完備2因子実験計画に対する, 表 10.2 と同様なモデル II 分散分析表は練習問題とし

て読者に残しておく.

(c) 完備3因子実験計画

完備3因子実験計画における分散成分の推定は, 8.6(c)節で与えた3次元行列標本に関する標本論からでてくる. ここではその節の記号を用いる. 完備3因子実験計画に関する応答確率変数の標本は8.6(c)節で述べた行列標本 $\{x_{\xi\eta\zeta}\}$ である. もちろん, ここで r 個の R 要因水準, s 個の C 要因水準, t 個の L 要因水準は無限母集団からの標本であるとする. このように, 標本 $\{x_{\xi\eta\zeta}\}$ の各要素 $x_{\xi\eta\zeta}$ は (8.6.30) で述べたように平均0, 共分散0の確率変数の和である. すなわち

(10.8.7)
$$x_{\xi\eta\zeta} = \mu + \varepsilon_{\xi\cdot\cdot} + \varepsilon_{\cdot\eta\cdot} + \varepsilon_{\cdot\cdot\zeta} + \varepsilon_{\xi\eta\cdot} + \varepsilon_{\xi\cdot\zeta} + \varepsilon_{\cdot\eta\zeta} + \varepsilon_{\xi\eta\zeta}$$

$x_{\xi\eta\zeta}$ の分散 σ^2 は (8.6.27) または (8.6.28) で述べたように, 7成分の和である. すなわち

(10.8.8)
$$\sigma^2 = \sigma^2_{\cdot 00} + \sigma^2_{0\cdot 0} + \sigma^2_{00\cdot} + \sigma^2_{\cdot\cdot 0} + \sigma^2_{\cdot 0\cdot} + \sigma^2_{0\cdot\cdot} + \sigma^2_{\cdot\cdot\cdot}$$

さて **10.1.1** (または **10.7.1**) を用いれば, 等式 (8.6.34) を分散成分に関して解くことによって, (8.6.31) で定義した $S_{\cdot 00}$, $S_{0\cdot 0}$, $S_{00\cdot}$, $S_{\cdot\cdot 0}$, $S_{\cdot 0\cdot}$, $S_{0\cdot\cdot}$, $S_{\cdot\cdot\cdot}$ から, 分散成分に対する推定量を得ることができる. すなわち

(10.8.9)
$$\begin{aligned}
\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot\cdot}) &= \frac{S_{\cdot\cdot\cdot}}{(r-1)(s-1)(t-1)}\\
\mathscr{E}^{-1}(\sigma^2_{0\cdot\cdot}) &= \frac{S_{0\cdot\cdot}}{r(s-1)(t-1)} - \frac{1}{r}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot\cdot})\\
\mathscr{E}^{-1}(\sigma^2_{\cdot 0\cdot}) &= \frac{S_{\cdot 0\cdot}}{s(r-1)(t-1)} - \frac{1}{s}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot\cdot})\\
\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot 0}) &= \frac{S_{\cdot\cdot 0}}{t(r-1)(s-1)} - \frac{1}{t}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot\cdot})\\
\mathscr{E}^{-1}(\sigma^2_{00\cdot}) &= \frac{S_{00\cdot}}{rs(t-1)} - \frac{1}{s}\mathscr{E}^{-1}(\sigma^2_{0\cdot\cdot}) - \frac{1}{r}\mathscr{E}^{-1}(\sigma^2_{\cdot 0\cdot}) - \frac{1}{rs}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot\cdot})\\
\mathscr{E}^{-1}(\sigma^2_{0\cdot 0}) &= \frac{S_{0\cdot 0}}{rt(s-1)} - \frac{1}{t}\mathscr{E}^{-1}(\sigma^2_{0\cdot\cdot}) - \frac{1}{r}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot 0}) - \frac{1}{rt}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot\cdot})\\
\mathscr{E}^{-1}(\sigma^2_{\cdot 00}) &= \frac{S_{\cdot 00}}{st(r-1)} - \frac{1}{t}\mathscr{E}^{-1}(\sigma^2_{\cdot 0\cdot}) - \frac{1}{s}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot 0}) - \frac{1}{st}\mathscr{E}^{-1}(\sigma^2_{\cdot\cdot\cdot})
\end{aligned}$$

μ に対する不偏推定量は $\bar{x}_{\cdot\cdot\cdot}$ であり, その分散は式 (8.6.35) で与えられる. もちろん $\sigma^2(\bar{x}_{\cdot\cdot\cdot})$ に対する不偏推定量は (8.6.35) の各分散成分を (10.8.9) におけるそれらの推定量で置き換えれば得られる.

完備3因子実験計画に関する表10.2と同様なモデルII分散分析表の作成は読者の練習問題としよう.

R 要因, C 要因, L 要因の各水準の母集団が有限である場合にも, 完備3因子実験計画における分散成分の推定はそのまま適用できることを指摘しておこう. これは読者に残しておくことにする.

(d) ラテン方格実験計画

8.6(d)節と10.6(c)節で指摘してきたように, ラテン方格実験計画は釣合い不完備3次元行列標本の特別な場合である. すなわち $s=t=r$, かつ $s'=t'=r'=1$ の場合である. ラテン方格実験計画に対する応答確率変数の標本 $\{x_{\xi\eta\zeta}\}^*$ は (ただし $\{x_{\xi\eta\zeta}\}^*$ は8.6(d)節で定義されたもの), この標本に含まれる各 $x_{\xi\eta\zeta}$ が (10.8.7) の形をしたものであり, その分散 σ^2 は, (10.8.8) で与えられた7つの成分からなっている. しかしながら, この標本に関する総平方和 S^*_T が (8.6.40) で定義された平方和 $S^*_{\cdot 00}$, $S^*_{0\cdot 0}$, $S^*_{00\cdot}$, S^*_E に分解できるならば, (8.6.41) で見たように (10.8.8) に含まれる次の成分は推定できる.

(10.8.10)
$$\sigma^2_E,\qquad \sigma^2_{\cdot 00},\qquad \sigma^2_{0\cdot 0},\qquad \sigma^2_{00\cdot}$$

ただし

$$\sigma^2_E = \sigma^2_{\cdot\cdot\cdot} + \sigma^2_{\cdot\cdot 0} + \sigma^2_{\cdot 0\cdot} + \sigma^2_{0\cdot\cdot}$$

言い換えれば, 行, 列, 層の各成分と, 残り4成分の和のみしか推定することができない. もちろん, これらの推定は (8.6.41) に **10.7.1** を適用することによって得られる. すなわち

(10.8.11)
$$\begin{aligned}
\mathscr{E}^{-1}(\sigma^2_{\cdot 00}) &= \frac{S^*_{\cdot 00}}{r-1} - \mathscr{E}^{-1}(\sigma^2_E)\\
\mathscr{E}^{-1}(\sigma^2_{0\cdot 0}) &= \frac{S^*_{0\cdot 0}}{r-1} - \mathscr{E}^{-1}(\sigma^2_E)\\
\mathscr{E}^{-1}(\sigma^2_{00\cdot}) &= \frac{S^*_{00\cdot}}{r-1} - \mathscr{E}^{-1}(\sigma^2_E)\\
\mathscr{E}^{-1}(\sigma^2_E) &= \frac{S^*_{\cdot 00} + S^*_{0\cdot 0} + S^*_{00\cdot}}{(2r-1)} - \frac{S^*_E}{(2r-1)(r-1)}.
\end{aligned}$$

ラテン方格実験計画に関する表10.2と同様なモデルII分散分析表は容易にできる. 練習問題として読者に残してこう.

10.9　層化母集団平均に対する線形推定量

(a)　定義と記号

有限母集団 π_N からの標本抽出に関する実際問題では，母集団 π_N は m 個の互いに素な層 $\pi_{N_1},\cdots,\pi_{N_m}$ に分解されていることがよくある．ただし $N_1+\cdots+N_m=N$ で π_{N_g} の平均と分散は $\mu_g,\ \sigma_g^2,\ g=1,\cdots,m$ である． π_N の平均と分散を $\mu,\ \sigma^2$ とすれば， μ と σ^2 は $N_g,\ \mu_g,\ \sigma_g^2,\ g=1,\cdots,m$ を用いて次のように表わすことができる．

(10.9.1) $$\mu=\sum_g p_g\mu_g$$

(10.9.2) $$\sigma^2=\sigma_W^2+\sigma_B^2.$$

ただし

(10.9.3) $$\sigma_W^2=\sum_g\left(\frac{N_g-1}{N-1}\right)\sigma_g^2,\qquad \sigma_B^2=\frac{N}{N-1}\sum_g p_g(\mu_g-\mu)^2$$

(10.9.4) $$p_g=\frac{N_g}{N},\qquad g=1,\cdots,m.$$

σ^2 は2つの成分，すなわち層内変動要因 σ_W^2 と層間変動要因 σ_B^2 の和であることに注意しよう．

$O_{n_1},\cdots,O_{n_m}$ が層 $\pi_{N_1},\cdots,\pi_{N_m}$ からの大きさ $n_1,\cdots,n_m$ のそれぞれ独立な標本のとき，$\bar{x}_1,\cdots,\bar{x}_m$ をそれぞれ $O_{n_1},\cdots,O_{n_m}$ の標本平均，$s_1^2,\cdots,s_m^2$ を標本分散とする． ひとまとめにとった層標本 $O_{n_1},\cdots,O_{n_m}$ (ただし $n_1+\cdots+n_m=n$) を π_N からの大きさ n の層化標本と呼ぶ．

ここで主な問題は， $N_g,\ \mu_g,\ \sigma_g^2,\ g=1,\cdots,m$ に関して与えられた種々の情報に対して各層からの標本平均に基づいて，μ に対する最適な線形推定量を求めることである．実際に，下記の2種類の条件の下でこの問題を考えよう．

(i)　層の大きさ $N_1,\cdots,N_m$ は既知である．

(ii)　層の大きさの他に，各層の分散 $\sigma_1^2,\cdots,\sigma_m^2$ が既知である．

母集団 π_N が層化されていない場合は， すでに10.2節で見たように， π_N からとった単一の無作為標本 O_n による μ の最小分散線形推定量は $\bar{x}$ であり，$\bar{x}$ の分散は $(1/n-1/N)\sigma^2$ であった．この結果は条件 (i) および (ii) に対して求められる推定量の分散を比較できるような“基準”を与えてくれる．

層化標本抽出法は政治，経営，工業関係の標本調査に広く用いられており，この種の問題に関する文献は多いが，ここでは数学的な序論を述べよう．それ以上のことを学びたい読者は Cochran (1953), Hansen, Hurwitz と Madow (1953), Stephan と McCarthy (1958), Sukhatme (1954), Yates (1949) の書物を参照されたい．

(b)　一般層化標本からの母集団平均に対する線形推定量

この場合には層の大きさ $N_1,\cdots,N_m$ は既知で， 層標本の大きさ $n_1,\cdots,n_m,\ n_i\geqslant 2$ は指定されているものと仮定する．$(x_{g1},\cdots,x_{gn_g})$ を標本 O_{n_g} の要素としよう．ただし $g=1,\cdots,m$． この層標本の組を π_N からの一般層化標本と呼ぶ．L をすべての標本要素の1次関数としよう．すなわち

(10.9.5) $$L=\sum_{g=1}^m\sum_{\xi=1}^{n_g}c_{g\xi}x_{g\xi}.$$

ここで L が μ に対する最小分散を持つような不偏推定量となるように $c_{g\xi}$ を定めたい．L が不偏であるためには，$c_{g\xi}$ は

$$\mathscr{E}(L)=\sum_{g=1}^m\sum_{\xi=1}^{n_g}c_{g\xi}\mu_g\equiv\sum_g p_g\mu_g$$

を満足するように選ばれねばならない．すなわち

(10.9.6) $$\sum_{\xi=1}^{n_g}c_{g\xi}=p_g,\qquad g=1,\cdots,m.$$

さらに，$\sigma^2(L)$ を最小にするように選ぶ必要がある．層標本は独立であるから次が得られる．

(10.9.7) $$\sigma^2(L)=\sum_{g=1}^m\left[\sigma_g^2\left(\sum_{\xi=1}^{n_g}c_{g\xi}^2\right)-\frac{\sigma_g^2}{N_g}\left(\sum_{\xi=1}^{n_g}c_{g\xi}\right)^2\right].$$

(10.9.6) を満足し，かつ (10.9.7) を最小にするような $c_{g\xi}$ の値は次のようになることが証明できる．

(10.9.8) $$c_{g\xi}=\frac{p_g}{n_g},\qquad \xi=1,\cdots,n_g,\quad g=1,\cdots,m.$$

これらの値を (10.9.5) に代入して得られる μ に対する一般層化標本の推定量を $\mathscr{E}_s^{-1}(\mu)$ で表わすと

(10.9.9) $$\mathscr{E}_s^{-1}(\mu)=p_1\bar{x}_1+\cdots+p_m\bar{x}_m.$$

$\mathscr{E}_s^{-1}(\mu)$ の分散は

(10.9.10)
$$\sigma^2(\mathscr{E}_s^{-1}(\mu))=\sum_g p_g^2\left(\frac{1}{n_g}-\frac{1}{N_g}\right)\sigma_g^2.$$

これから

(10.9.11)
$$\sum_g p_g^2\left(\frac{1}{n_g}-\frac{1}{N_g}\right)s_g^2$$

が $\sigma^2(\mathscr{E}_s^{-1}(\mu))$ に対する不偏推定量であることがわかる．以上をまとめると，

10.9.1　π_N を大きさが既知の互いに素な層 $\pi_{N_1},\cdots,\pi_{N_m}$ を持つ有限母集団とする．また π_{N_g} の平均と分散を μ_g, σ_g^2, $g=1,\cdots,m$ とする．$O_{n_1},\cdots,O_{n_m}$ を π_N からの一般層化標本として，O_{n_g} の標本平均，標本分散をそれぞれ $\bar{x}_g$, s_g^2, $g=1,\cdots,m$ とする．このとき，π_N の平均に対して (10.9.9) で与えられる $\mathscr{E}_s^{-1}(\mu)$ は μ に対する最小分散線形不偏推定量である．$\mathscr{E}_s^{-1}(\mu)$ の分散は (10.9.10) で与えられ，この分散に対する不偏推定量は (10.9.11) で与えられる．

前述の議論では層標本の大きさ $n_1,\cdots,n_m$, $n_i\geqq 2$ は固定されていたことに注目しよう．さらに，$\mathscr{E}_s^{-1}(\mu)$ は層標本の大きさの任意の選び方に対して，π_N の平均に対する不偏推定量となる．

層標本の大きさ $n_1,\cdots,n_m$ が層の大きさに比例して選ばれるような特別の場合，すなわち $n_g=p_gn$, $g=1,\cdots,m$ のとき π_N から大きさ n の比例層化標本が得られる（実際問題では，p_gn はほとんど整数に近い）．n_g, $g=1,\cdots,m$ をこのように選んだとき，$\mathscr{E}_s^{-1}(\mu)$ を $\mathscr{E}_{sp}^{-1}(\mu)$ と表わす．$\mathscr{E}_{sp}^{-1}(\mu)$ の分散は

(10.9.10 a)
$$\sigma^2(\mathscr{E}_{sp}^{-1}(\mu))=\left(\frac{1}{n}-\frac{1}{N}\right)\sum_g p_g\sigma_g^2$$

であるから

(10.9.11 a)
$$\left(\frac{1}{n}-\frac{1}{N}\right)\sum_g p_gs_g^2$$

は $\sigma^2(\mathscr{E}_{sp}^{-1}(\mu))$ に対する不偏推定量である．

π_N の平均 μ を推定する場合，π_N からの大きさ n の単純無作為標本の平均 $\bar{x}$ に対する $\mathscr{E}_{sp}^{-1}(\mu)$ の効率は比

(10.9.12)
$$\mathrm{eff}(\mathscr{E}_{sp}^{-1}(\mu))=\frac{\sigma^2(\bar{x})}{\sigma^2(\mathscr{E}_{sp}^{-1}(\mu))}=1+K$$

で定義される．ただし

(10.9.13)
$$K=\frac{\sigma_B^2}{\sigma_W^2}+O\left(\frac{1}{N}\right).$$

π_N の平均を線形推定する場合，比 σ_B^2/σ_W^2 より，比例層化抽出は無作為抽出よりすぐれていることが (10.9.12) から明らかである．$\sigma_B^2=0$, すなわち，すべての層が等平均を持つときは，比例層化抽出は単純無作為抽出より有利にはならない．

まとめると，

10.9.2　比例層化抽出 ($n_g=p_gn$, $g=1,\cdots,m$ として選ぶ) の結果得られる線形推定量 $\mathscr{E}_{sp}^{-1}(\mu)$ は分散 (10.9.10 a) を持ち，その不偏推定量は (10.9.11 a) となる．さらに，π_N の平均に対する推定量として π_N からの大きさ n の無作為標本の平均 $\bar{x}$ に対する $\mathscr{E}_{sp}^{-1}(\mu)$ の効率は (10.9.12) で与えられる．

(c)　母集団の層の大きさ，分散が既知の場合の母集団平均に対する最小分散線形推定量

$N_1,\cdots,N_m$ ばかりでなく，$\sigma_1^2,\cdots,\sigma_m^2$ の値もわかっているときは，比例層化抽出より層標本の大きさを“より良く”選択できる．ここでの問題は (10.9.10) で与えたように $n_1+\cdots+n_m=n$ なる条件に従って，$\sigma^2(\mathscr{E}_s^{-1}(\mu))$ を最小化することである．いま $n_1,\cdots,n_m$ を連続的に変えてもよいことにすれば，与えられた条件を満足する $n_1,\cdots,n_m$ の値は次のようになることが証明できる．

(10.9.14)
$$n_g=b_gn,\qquad g=1,\cdots,m$$

ただし

(10.9.15)
$$b_g=p_g\sigma_g\Big/\sum_g p_g\sigma_g.$$

$n_1,\cdots,n_m$ のこの最適選択に対して，$\mathscr{E}_s^{-1}(\mu)$ を $\mathscr{E}_{s0}^{-1}(\mu)$ で表わす．このときの (10.9.10) を簡単にすると

(10.9.10 b)
$$\sigma^2(\mathscr{E}_{s0}^{-1}(\mu))=\frac{1}{n}\left(\sum_g p_g\sigma_g\right)^2-\frac{1}{N}\sum_g p_g\sigma_g^2.$$

（実際問題では，もちろん各 n_g に対して，nb_g に最も近い正整数を選べばよい．）

(10.9.16)
$$\sigma^2(\mathscr{E}_{s0}^{-1}(\mu))\leqq\sigma^2(\mathscr{E}_{sp}^{-1}(\mu))$$

が成り立ち，等号が成立するのは $\sigma_1=\cdots=\sigma_m$ のときかつそのときに限る．なぜなら，

(10.9.10) と (10.9.10 a) を比較すれば，(10.9.16) が成り立つ必要十分条件は

$$\left(\sum_g p_g\sigma_g\right)^2 \leqslant \sum_g p_g\sigma_g^2.$$

すなわち

(10.9.17)
$$\left(\sum_g p_g\sigma_g\right)^2 \leqslant \left(\sum_g p_g\right)\left(\sum_g p_g\sigma_g^2\right)$$

である．これは $\sum_g p_g = 1$ で，p_g はすべて正であるからである．

しかし，(10.9.17) は有名なシュワルツの不等式の特別な場合であり，等式が成り立つのは $\sigma_1 = \cdots = \sigma_m$ のときかつそのときに限る．

π_N からの大きさ n の層化抽出は，n_g が (10.9.14) を満たすとき，π_N からの大きさ n の**最適層化抽出**と呼ばれる．

まとめると，Neyman (1934) による次の結果が得られる．

10.9.3　π_N を有限母集団で，大きさ $N_1, \cdots, N_m$ の互いに素な層 $\pi_{N_1}, \cdots, \pi_{N_m}$ を持ち，その分散 $\sigma_1^2, \cdots, \sigma_m^2$ は既知とする．このとき，π_N の平均に対する最小分散線形推定量 $\mathscr{E}_{s0}^{-1}(\mu)$ を与える π_N からの大きさ n の層化標本は，その層と標本の大きさが (10.9.14) で与えられた n_g になっている．$\mathscr{E}_{s0}^{-1}(\mu)$ の分散は (10.9.10 b) で与えられる．$\mathscr{E}_{s0}^{-1}(\mu)$ および $\mathscr{E}_{sp}^{-1}(\mu)$ は層の分散がすべて等しいときかつそのときに限って，π_N の平均を推定するにあたって等効率を持つ．

(d)　総費用が固定している場合の母集団平均に対する最小分散線形推定量

実際問題では，種々の層から標本抽出に要する費用はかなりの変動があることが多い．このような場合，総費用を固定しておいて，(10.9.10) での $\sigma^2(\mathscr{E}_s^{-1}(\mu))$ を最小にするように，いくつかの層からの標本の大きさを選択することが望ましい．正確にいえば，c_g を π_{N_g} からの大きさ n_g の標本中の各要素の値を求める費用とする．C を総費用とすれば

(10.9.18)
$$C = c_1n_1 + \cdots + c_mn_m.$$

(10.9.18) を条件として，$\sigma^2(\mathscr{E}_s^{-1}(\mu))$ が最小となるのは

(10.9.19)
$$n_g = CB_g, \qquad g = 1, \cdots, m.$$

ただし

(10.9.20)
$$B_g = \frac{\sqrt{c_g}\, p_g\sigma_g}{c_g \sum_g (\sqrt{c_g}\, p_g\sigma_g)}$$

のときであることは容易にわかる．この方法では，費用 $c_1, \cdots, c_m$ ばかりでなく，層の大きさ $N_1, \cdots, N_m$ および層の分散 $\sigma_1^2, \cdots, \sigma_m^2$ が利用できる（既知）ことを仮定していることに注意しておこう．

(10.9.19) で示したように n_g を選んだとき，$\mathscr{E}_s^{-1}(\mu)$ を $\mathscr{E}_{sc}^{-1}(\mu)$ で表わす．このとき

(10.9.21)
$$\sigma^2(\mathscr{E}_{sc}^{-1}(\mu)) = \frac{1}{C}\left(\sum_g \sqrt{c_g}\, p_g\sigma_g\right)^2 - \frac{1}{N}\sum_g p_g\sigma_g^2.$$

(e)　無限母集団からの層化抽出への拡張

10.9.1, **10.9.2**, **10.9.3** の結果はそのまま無限母集団からの層化抽出の場合に拡張できる．ただし，$N_1, \cdots, N_m \longrightarrow \infty$ のとき，p_g, σ_g^2, $g = 1, 2, \cdots$ は正の値に収束し，μ_g は有限であるとする．この拡張は練習問題として読者に残しておこう．

10.10　2段抽出における層化母集団の平均に対する線形推定

(a)　2 段 抽 出

π_N が非常に多くの層を持っている場合には，その各層から標本を抽出するのは経済的に不可能と思われる．このようなとき，**2段抽出法**を用いるとよい．すなわち，まず第1に，特定の確率で若干の層を無作為に選び出し，次にこれらの各層から単純無作為抽出によって要素を引き出す．標本抽出の目的は平均値 μ に対する推定量，この推定量の分散およびこの分散に対する推定量を得ることである．

簡単のために，各層の大きさが整数 ν の倍数になっている場合を考えよう．すなわち

(10.10.1)
$$N_g = M_g\nu, \qquad g = 1, \cdots, m.$$

ν を**抽出単位**としよう．このとき，g 番目の層は M_g 個の標本抽出単位からなるとみなしてよい．したがって π_N における標本抽出単位の合計数は，$M_1 + \cdots + M_m = M$ である．ゆえに

(10.10.2)
$$N = M\nu$$

である．

さて，取り出される標本抽出単位の合計数が u であるという条件の下で， 母集団の各層から重複を許さないで取り出される抽出単位の数を定めるための確率過程を考えよう．これは Wilks (1960 b) による．この過程では，M 個の抽出単位はおのおの等確率で取り出されると仮定する．（これは乱数を用いれば実現できる．）δ_g を π_{N_g}，$g=1,\cdots,m$ から取り出される抽出単位の数を表わす確率変数とする．このとき，$(\delta_1,\cdots,\delta_m)$ は $(m-1)$ 次元確率変数で，$(m-1)$ 次元超幾何分布を持ち，p.f. は

(10.10.3)
$$p(\delta_1,\cdots,\delta_m)=\frac{\binom{M_1}{\delta_1}\cdots\binom{M_m}{\delta_m}}{\binom{M}{u}}$$

となる．ただし $\delta_1+\cdots+\delta_m=u$.

したがって，標本抽出の第1段の結果は，本質的には大きさ $\delta_g v$ の標本が π_{N_g}，$g=1,\cdots,m$ から抽出されることになる．

標本抽出の第2段では，大きさ $\delta_1 v,\cdots,\delta_m v$ の標本をそれぞれ $\pi_{N_1},\cdots,\pi_{N_m}$ から抽出する．この過程は π_N からの大きさ uv の2段抽出法を与える．これは抽出する層を無作為に選んだ層化抽出とみなすことができる． 実際には，層の数が多いときには多くの δ が 0 となっている．すなわちまったく抽出されない層が沢山ある．抽出される層の合計数は 0 でない δ の個数である．

(b) 母集団平均に対する線形推定量

$\mathscr{E}_2^{-1}(\mu)$ を次のような μ に対する線形推定量とする．

$$\mathscr{E}_2^{-1}(\mu)=\frac{(\delta_1 v)\bar{x}_1+\cdots+(\delta_m v)\bar{x}_m}{uv}.$$

すなわち

(10.10.4)
$$\mathscr{E}_2^{-1}(\mu)=\frac{\delta_1\bar{x}_1+\cdots+\delta_m\bar{x}_m}{u}.$$

すると

(10.10.5)
$$\mathscr{E}(\mathscr{E}_2^{-1}(\mu))=\sum_g\mathscr{E}\left(\frac{\delta_g\bar{x}_g}{u}\right)$$

が得られる．3.7.2 で述べた繰り返し平均値を用いると，$g=1,\cdots,m$ に対して次が得られる．

(10.10.6)
$$\mathscr{E}\left(\frac{\delta_g\bar{x}_g}{u}\right)=\frac{1}{u}\mathscr{E}_{(\delta_g)}[\delta_g\cdot\mathscr{E}_{(\bar{x}_g)}(\bar{x}_g|\delta_g)]$$

ただし $\mathscr{E}_{(\bar{x}_g)}(\bar{x}_g|\delta_g)$ は条件つき確率変数 $\bar{x}_g|\delta_g$ の平均値である．一方，$\mathscr{E}_{(\delta_g)}[\quad]$ は確率変数 δ_g に関する非条件つきの平均値を表わしている．しかし

$$\mathscr{E}_{(\bar{x}_g)}(\bar{x}_g|\delta_g)=\mu_g$$

であり，また (6.1.8) から

$$\mathscr{E}_{(\delta_g)}(\delta_g)=p_g u.$$

したがって

(10.10.7)
$$\mathscr{E}(\mathscr{E}_2^{-1}(\mu))=\sum_g p_g\mu_g=\mu.$$

これは $\mathscr{E}_2^{-1}(\mu)$ が π_N の平均値 μ に対する不偏推定量になっていることを示している．

$\mathscr{E}_2^{-1}(\mu)$ の分散は

(10.10.8)
$$\sigma^2(\mathscr{E}_2^{-1}(\mu))=\mathscr{E}\left[\sum_g\left(\frac{\delta_g\bar{x}_g}{u}-p_g\mu_g\right)\right]^2$$
$$=\mathscr{E}\left[\sum_g\frac{\delta_g}{u}(\bar{x}_g-\mu_g)+\sum_g\left(\frac{\delta_g}{u}-p_g\right)\mu_g\right]^2$$

で定義される．

この平方を計算して，繰り返し平均値をとり，簡単にすれば

(10.10.9)
$$\sigma^2(\mathscr{E}_2^{-1}(\mu))=\frac{N-uv}{u(N-v)}\left[\frac{N-1}{N}\left(\frac{\sigma_W^2}{v}+\sigma_B^2\right)-\frac{v-1}{Nv}\sum_g\sigma_g^2\right].$$

(10.9.3) を参照すれば，$\sigma^2(\mathscr{E}_2^{-1}(\mu))$ は次の3つの量の1次関数であることがわかる．

(10.10.10)
$$\sum_g\sigma_g^2,\qquad\sum_g p_g\sigma_g^2,\qquad\sum_g p_g(\mu_g-\mu)^2.$$

G_1，G_2，G_3 を次のように定義された2段標本要素の関数としよう．ただし，$v\geqslant 2$，

(10.10.11)
$$G_1=\sum_g\delta_g s_g^2,\qquad G_2=\sum_g\frac{\delta_g}{N_g}s_g^2,\qquad G_3=\sum_g\delta_g(\bar{x}_g-\mathscr{E}_2^{-1}(\mu))^2$$

$\mathscr{E}(G_1)$，$\mathscr{E}(G_2)$，$\mathscr{E}(G_3)$ が (10.10.10) における3つの量の1次関数であって，$\sigma^2(\mathscr{E}_2^{-1}(\mu))$ に対する不偏推定量となるような G_1，G_2，G_3 の1次関数が存在することはただちに証明できる．実際，この不偏推定量は (10.10.9) において，σ_W^2，σ_B^2，$\sum_g\sigma_g^2$ をそれぞれ次の不偏推定量で置き換えれば得られる．

(10.10.12)
$$\mathscr{E}^{-1}(\sigma_W^2)=\frac{N}{u(N-1)}(G_1-G_2)$$
$$\mathscr{E}^{-1}(\sigma_B^2)=\frac{Nv}{uv(N-1)+N}\left[\frac{N-2}{uv(N-1)}(G_1-G_2)-\frac{Nu+2v}{u^2v}G_2+G_3\right]$$

$$\mathcal{E}^{-1}\left(\sum_g \sigma_g^2\right) = \frac{N}{u} G_2.$$

以上の結果をまとめると，

10.10.1　π_N を互いに素な層 $\pi_{N_1}, \cdots, \pi_{N_m}$ からなる有限母集団とする．ただし $N_g = M_g v$, $g=1,\cdots,m$, また v は与えられた整数（標本抽出単位）．u 個の抽出単位からなる2段抽出で π_N から標本抽出する．このとき $\mathcal{E}_2^{-1}(\mu)$ は π_N の平均の線形不偏推定量となる．$\mathcal{E}_2^{-1}(\mu)$ の分散は (10.10.9) で与えられる．一方 (10.10.9) の右辺の σ_W^2, σ_B^2, $\sum_g \sigma_g^2$ を (10.10.12) の推定量で置き換えた統計量は，$v \geqq 2$ に対する $\sigma^2(\mathcal{E}_2^{-1}(\mu))$ の不偏推定量となる．

(c) 無限母集団の場合

層 $\pi_1, \cdots, \pi_m$ がそれぞれ $g=1,\cdots,m$ に対して，$N_g \longrightarrow \infty$ のとき，p_g と σ_g^2 が正で，μ_g が有限であるような無限層化母集団の場合には，確率変数 $(\delta_1, \cdots, \delta_m)$ は多項 p.f.

(10.10.13)
$$p(\delta_1, \cdots, \delta_m) = \frac{u\,!}{\delta_1! \cdots \delta_m!} p_1^{\delta_1} \cdots p_m^{\delta_m}$$

を持つ．ただし $\delta_1 + \cdots + \delta_m = u$.

この場合，$\mathcal{E}_2^{-1}(\mu)$ の分散は

(10.10.14)
$$\sigma^2(\mathcal{E}^{-1}(\mu)) = \frac{\sigma_{W'}^2}{uv} + \frac{\sigma_{B'}^2}{u}.$$

ただし

$$\sigma_{W'}^2 = \sum_g p_g \sigma_g^2, \qquad \sigma_{B'}^2 = \sum_g p_g (\mu_g - \mu)^2.$$

一方

(10.10.15)
$$\frac{n+v+1}{un(n+1)} G_1 + \frac{v^2}{n(n+1)} G_3 - \frac{v}{un(n+1)} \sum_g \frac{\delta_g}{p_g} s_g^2$$

は $n = uv$ のとき，$\sigma^2(\mathcal{E}_2^{-1}(\mu))$ の不偏推定量である．ただし $\sigma_{B'}^2 = \sigma_B^2 + O(1/N)$, $\sigma_{W'}^2 = \sigma_W^2 + O(1/N)$ が成り立つ．

(d) 総費用が固定されている場合の母集団平均に対する最小分散線形推定量

標本の大きさ uv が固定されているとしよう．$uv = n$ と置けば，$\sigma^2(\mathcal{E}_2^{-1}(\mu))$ は次のように書ける．

(10.10.16)
$$\sigma^2(\mathcal{E}_2^{-1}(\mu)) = \frac{N-n}{n}\left(\frac{A+NB}{N-v} - B\right).$$

ただし

(10.10.17)
$$A = \sigma_{W'}^2, \qquad B = \sigma_{B'}^2 - \frac{1}{N}\sum_g \sigma_g^2.$$

(10.10.16) から，v は正整数でなければならないから，もちろん $A + NB \geqq 0$ で n 個の層があるならば，$\sigma^2(\mathcal{E}_2^{-1}(\mu))$ を最小にする v の値は $v = 1$ であることは明らかである．このとき，π_N からの2段抽出は π_N からの大きさ n の単純無作為標本に帰着する．この場合には，(10.10.9) で与えられた $\mathcal{E}_2^{-1}(\mu)$ の分散は，$(1/n - 1/N)\sigma^2$ となる．ただし σ^2 は (10.9.2) で与えられている．

この状態からみれば，与えられた状態における u と v の値の選択は $\sigma^2(\mathcal{E}_2^{-1}(\mu))$ を最小にする基準ではできない．その理由はこの基準では $v=1$ となるはずだからである．実際問題では，u, v の選択基準は総費用が固定された場合に，$\sigma^2(\mathcal{E}_2^{-1}(\mu))$ を最小にするように u と v の値の組を求めることである．この場合，最も簡単な方法は抽出単位あたりの費用 C_1，標本要素あたりの費用 C_2 を仮定することである．このときの総費用 C は

(10.10.18)
$$C = C_1 u + C_2 uv$$

になる．(10.10.9) で与えた $\sigma^2(\mathcal{E}_2^{-1}(\mu))$ が制約 (10.10.18) のもとで最小になるならば，u, v を連続的に動かすと最小になる v の値（$\bar{v}$ とすれば）は，

(10.10.19)
$$\bar{v} = \frac{N}{1 + \sqrt{\left(1 + \dfrac{H}{G} N\right)\left(1 + \dfrac{B}{A} N\right)}}$$

で与えられる．ただし

(10.10.20)
$$G = N\frac{C_1}{C}, \qquad H = N\frac{C_2}{C} - 1.$$

また A と B は (10.10.17) で与えられている．これに対応する u の値 $\bar{u}$ は，もちろん (10.10.18) の v に $\bar{v}$ を代入して u について解けば得られる．これらの条件下での推定量 $\mathcal{E}_2^{-1}(\mu)$ を $\mathcal{E}_{2c}^{-1}(\mu)$ で表わせば，$\sigma^2(\mathcal{E}_{2c}^{-1}(\mu))$ の値は (10.10.9) の u と v を $\bar{u}$ と $\bar{v}$ で置き換えたもので与えられる．同様に，$\bar{u}$ と $\bar{v}$ の値を (10.10.12) の u, v に代入すれば，$\sigma^2(\mathcal{E}_{2c}^{-1}(\mu))$ の不偏推定量が得られる．N が大きい場合には，

$$(10.10.21)\qquad \bar{v}=\frac{\sigma_{W'}}{\sigma_{B'}}\sqrt{\frac{C_1}{C_2}}+O\left(\frac{1}{N}\right),\qquad \bar{u}=\frac{C}{C_1+\frac{\sigma_{W'}}{\sigma_{B'}}\sqrt{C_1C_2}}+O\left(\frac{1}{N}\right)$$

になる．

問　　題

10.1　$\bar{x}$ を大きさ N の有限母集団からの大きさ n の標本の標本平均とした場合に，**10.2.1** を証明せよ．

10.2　$(x_1,\cdots,x_n)$ が n 次元確率変数で，各 x の平均が μ，分散が σ^2，共分散が $\rho\sigma^2$ となる分布を持つと仮定する．このとき，μ に対する最小分散線形推定量は n 個の確率変数 $x_1,\cdots,x_n$ の平均 $\bar{x}$ であることを示せ．

10.3　$\bar{x}$, $\bar{x}'$ をそれぞれ平均 μ, μ' および分散 σ^2, σ'^2 を持つ分布からの大きさ n, n' の独立な標本の標本平均とするとき，$\bar{x}-\bar{x}'$ は $\mu-\mu'$ に対する最小分散線形推定量であることを示せ．

10.4　$T_1,\cdots,T_k$ がパラメータ θ の不偏推定量で，（正則な）共分散行列 $\|\sigma_{ij}\|$ と逆行列 $\|\sigma^{ij}\|$ を持ち，$l_1,\cdots,l_k$ を和が1になる定数とすれば，$\sum_{i=1}^{k} l_iT_i$ なる形をした θ の最小分散推定量が $l_i=\sum_j \sigma^{ij}/\sum_{i,j}\sigma^{ij}$, $i=1,\cdots,k$ のとき求まることを示せ．また推定量の分散は $1/\sum_{i,j}\sigma^{ij}$ になることを示せ．

10.5　**10.2.2** の結論は $(x_1,\cdots,x_n)$ を n 次元確率変数で，その c.d.f. $F(x_1,\cdots,x_n)$ が $x_1,\cdots,x_n$ の対称関数であると仮定した場合にも正しいことを示せ．

10.6　n_1, $\bar{x}_1$, s_1^2 を分布 $N(\mu_1,\sigma^2)$ からの標本の大きさ，平均，分散とし，また，n_2, $\bar{x}_2$, s_2^2 を $N(\mu_2,\sigma^2)$ からの独立な標本に関する同様な統計量とするとき，次を証明せよ．

$$(\bar{x}_1-\bar{x}_2)\pm t_{n_1+n_2-2,\gamma}s\sqrt{1/n_1+1/n_2}$$

は $\mu_1-\mu_2$ に対する $100\gamma\%$ 信頼区間である．ただし

$$s^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2},\qquad t_{n_1+n_2-2,\gamma}$$

は (10.2.10) で定義されている．

10.7　**10.3.2** を証明せよ．

10.8　**10.3.1** において $(x_{11},\cdots,x_{1n}),\cdots,(x_{k1},\cdots,x_{kn})$ を互いに直交する単位ベクトル（定数）として選べば，$\beta_1,\cdots,\beta_k$ に対する最小分散線形推定量 $\sum_\xi x_{i\xi}y_\xi$, $i=1,\cdots,k$ は分散がすべて σ^2 に等しく，共分散はすべて0であることを示せ．

10.9　**10.3.4** を参照して，楕円体

$$\sum_{i,j=1}^{k_1} a^*_{ij}(\beta_i-b_i)(\beta_j-b_j)=\frac{k_1S_1\mathscr{F}_{k_1,n-k,\gamma}}{(n-k)}$$

の内部は $(\beta_1,\cdots,\beta_{k_1})$, $k_1\leqq k$ に対する $100\gamma\%$ 信頼領域であることを示せ．ここに

$$\|a^*_{ij}\|=\left\|\begin{matrix} a^{11} & \cdots & a^{1k_1}\\ \cdot & & \cdot\\ \cdot & & \cdot\\ \cdot & & \cdot\\ a^{k_11} & \cdots & a^{k_1k_1}\end{matrix}\right\|^{-1}.$$

また $\mathscr{F}_{k_1,n-k,\gamma}$ は

$$\int_0^{\mathscr{F}_{k_1,n-k,\gamma}} dF_{k_1,n-k}(\mathscr{F})=\gamma$$

を満たし，$dF_{k_1,n-k}(\mathscr{F})$ はスネディッカー分布 $S(k_1,n-k)$ の p.e. とする．

10.10　R_n における球が中心 $(a_1,\cdots,a_n)$ と半径 r を持つとき，等式

$$\sum_{\xi=1}^{n} b_\xi x_\xi=k$$

を持つ超平面に平行でその球に接する2つの超平面（R_{n-1} 空間）の等式は次になることを示せ．

$$\sum_{\xi=1}^{n} b_\xi(x_\xi-a_\xi)=\pm r\sqrt{\sum_{\xi=1}^{n} b_\xi^2}.$$

10.11　10.5 節を参照して，$c_1,\cdots,c_k$ のすべては0でないようなあらゆる可能な実ベクトル $(c_1,\cdots,c_k)$ に対して，不等式

$$\sum_i c_ib_i-\delta\sqrt{\sum_{i,j}a_{ij}c_ic_j}<\sum_i c_i\beta_i<\sum_i c_ib_i+\delta\sqrt{\sum_{i,j}a_{ij}c_ic_j}$$

が同時に満たされる確率は γ であることを示せ．ただし b_i, β_i, S_1, $\mathscr{F}_{k,n-k,\gamma}$ は **10.3.4** と (10.4.4) で定義してあり，

$$\delta^2=\frac{kS_1}{(n-k)}\cdot\mathscr{F}_{k,n-k,\gamma}$$

とする．

10.12　$(u_1,\cdots,u_k)$ が分布 $N(\{\mu_i\};\|\sigma_{ij}\|)$ を持つと仮定する．ただし $(\mu_1,\cdots,\mu_k)$ は未知であるが $\|\sigma_{ij}\|$ は既知とする．$\chi^2_{k,\gamma}$ を自由度 k のカイ2乗分布の $100\gamma\%$ 点としよう．このとき，$c_1,\cdots,c_k$ のすべては0でないようなあらゆる可能な実ベクトル $(c_1,\cdots,c_k)$ に対して，不等式

$$\sum_i c_iu_i-\chi_{k,\gamma}\sqrt{\sum_{i,j}\sigma_{ij}c_ic_j}<\sum_i c_i\mu_i<\sum_i c_iu_i+\chi_{k,\gamma}\sqrt{\sum_{i,j}\sigma_{ij}c_ic_j}$$

が同時に満たされる確率は γ であることを示せ．

10.13　式 (10.6.5), (10.6.6) を確かめよ．

10.14　問題 8.20 において，μ に対する $100\gamma\%$ 信頼区間を求める方法を示せ．

10.15　(10.6.15), (10.6.16) を確かめよ．

10.16　完備3因子実験計画に対する **10.6.1**, **10.6.2** に相当する事柄を公式化し，これに対するモデルI分散分析表を作成せよ．

10.17　ラテン方格実験計画に対する **10.6.1**, **10.6.2** に相当する事柄を公式化し，これに対するモデルI分散分析表を作成せよ．

10.18　10.8(b)節で述べた不完備釣合い2因子実験計画に対するモデルII分散分析表を作成せよ．

10.19　不完備釣合い2因子実験計画に対する分散成分の推定量 (10.8.6) は $s'=s$, $r'=r$ のとき，(10.8.4) における完備2因子計画に対する推定量となることを確かめよ．

10.20　N_1 個の R 要因の水準，N_2 個の C 要因の水準からなる有限母集団の場合について，(10.8.4) に相当する推定量と，表10.2に相当する表を求めよ．

10.21　完備3因子実験計画に対するモデルII分散分析表を作成せよ．

10.22　ラテン方格実験計画に対するモデルII分散分析表を作成せよ．

10.23　完備2因子モデルI実験計画において，$\mu_{\cdot\eta}=0$, $\eta=1,\cdots,s$ ならば，その実験計画は完備1因子実験計画の s 回の反復として表わせる．この場合 $S^*_{\cdot\cdot}/[r(s-1)]$ は σ^2 の不偏推定量であることを示せ．ただし $S^*_{\cdot\cdot}=S_{\cdot\cdot}+S_{0\cdot}(0)$ である．**10.6.2** のように分布の正規性を仮定すれば，この場合の μ, $\mu_{\xi\cdot}$ に対する $100\gamma\%$ 信頼区間は次のようになることを示せ．

$$m\pm t_{r(s-1),\gamma}\sqrt{S^*_{\cdot\cdot}/[r^2s(s-1)]}$$
$$m_{\xi\cdot}\pm t_{r(s-1),\gamma}\sqrt{S^*_{\cdot\cdot}(r-1)/[r^2s(s-1)]}.$$

また

$$\sum_{\xi=1}^{r}(\mu_{\xi\cdot}-m_{\xi\cdot})^2=\frac{s(r-1)}{r(s-1)}S^*_{\cdot\cdot}\cdot\mathscr{F}_{(r-1),r(s-1),\gamma}$$

は条件 $\sum_{\xi=1}^{r}(\mu_{\xi\cdot}-m_{\xi\cdot})=0$ の下で，$(\mu_{1\cdot},\cdots,\mu_{r\cdot})$ に対する $100\gamma\%$ 信頼域を与える $(r-1)$ 次元の球形領域であることを示せ．

10.24　完備3因子モデルI実験計画において，$\mu_{\cdot\cdot\zeta}$, $\mu_{\cdot\eta\zeta}$, $\mu_{\xi\cdot\zeta}$ がすべて0ならば，r 個の行と s 個の列の完備2因子実験の t 回の反復が得られる．この場合，$S^*_{\cdot\cdot\cdot}/[rs(t-1)]$ は σ^2 に対する不偏推定量であることを示せ．ただし $S^*_{\cdot\cdot\cdot}=S_{\cdot\cdot\cdot}+S_{0\cdot\cdot}(0)+S_{\cdot0\cdot}(0)+S_{00\cdot}(0)$ である．また

$$\frac{1}{\sigma^2}S^*_{\cdot\cdot\cdot},\qquad\frac{1}{\sigma^2}S_{\cdot\cdot0}(\mu_{\xi\eta\cdot}),\qquad\frac{1}{\sigma^2}S_{\cdot00}(\mu_{\xi\cdot\cdot}),\qquad\frac{1}{\sigma^2}S_{0\cdot0}(\mu_{\cdot\eta\cdot})$$

はそれぞれ自由度 $rs(t-1)$, $(r-1)(s-1)$, $(r-1)$, $(s-1)$ を持つカイ2乗分布に従って独立に分布していることを示せ．それぞれ，(i) すべての $\mu_{\xi\eta\cdot}$ が0である，(ii) すべての $\mu_{\xi\cdot\cdot}$ が0である，(iii) すべての $\mu_{\cdot\eta\cdot}$ が0である，という仮説に対応するモデルI分散分析表を作成せよ．

10.25　10.6(b)節を参照して，信頼区間

$$(m_{\xi\cdot}-m_{\xi'\cdot})\pm\delta\sqrt{2(r-2)/rs},\qquad\xi\neq\xi'=1,\cdots,r$$

がすべて，それぞれ，$(\mu_{\xi\cdot}-\mu_{\xi'\cdot})$ を含むような確率は γ を越えることを示せ．ただし

$$\delta^2=\frac{S_{\cdot\cdot}}{(s-1)}\mathscr{F}_{(r-1),(r-1)(s-1),\gamma}$$

であり，$\mathscr{F}_{(r-1),(r-1)(s-1),\gamma}$ はスネディッカー分布 $S((r-1),(r-1)(s-1))$ の $100\gamma\%$ である．

10.26　$\bar{x}_1,\cdots,\bar{x}_k$, $s_1^2,\cdots,s_k^2$ をそれぞれ $N(\mu_1,\sigma^2),\cdots,N(\mu_k,\sigma^2)$ からの大きさ $n_1,\cdots,n_k$ の独立な標本の標本平均と標本分散とする．このとき，$\frac{1}{2}k(k-1)$ 個のすべての信頼区間 $(\bar{x}_i-\bar{x}_j)\pm\delta\sqrt{1/n_i+1/n_j}$ がそれぞれ $(\mu_i-\mu_j)$, $i>j=1,\cdots,k$ を含む確率は γ を越えることを示せ．ただし

$$\delta^2=\frac{k}{n-k}[(n_1-1)s_1^2+\cdots+(n_k-1)s_k^2]\mathscr{F}_{k,n-k,\gamma}$$

であり，$\mathscr{F}_{k,n-k,\gamma}$ はスネディッカー分布 $S(k,n-k)$ の $100\gamma\%$ 点である．

10.27　10.5(c)節と10.6(a)節の記号と定義を参照して不等式

$$(\bar{x}_{\xi\cdot}-\bar{x}_{\xi'\cdot})-D<(\mu_{\xi\cdot}-\mu_{\xi'\cdot})<(\bar{x}_{\xi\cdot}-\bar{x}_{\xi'\cdot})+D$$

がすべての $\xi\neq\xi'=1,\cdots,r-1$ に対して同時に成り立つ確率は γ を越えることを示せ．ただし

$$D=R_{r-2,(r-1)(s-1),\gamma}\sqrt{\frac{(r-2)S_{\cdot\cdot}}{rs(r-1)(s-1)}}.$$

10.28　10.5(c)節と10.6(b)節を参照して，固定された η に対して不等式

$$(m_{\xi\eta\cdot}-m_{\xi'\eta\cdot})-D<(\mu_{\xi\eta\cdot}-\mu_{\xi'\eta\cdot})<(m_{\xi\eta\cdot}-m_{\xi'\eta\cdot})+D$$

が $\xi\neq\xi'=1,\cdots,r-1$ のすべての選び方に対して同時に成り立つ確率は γ を越えることを示せ．ただし

$$D=R_{r-2,(r-1)(s-1)(t-1),\gamma}\sqrt{\frac{(r-2)S_{\cdot\cdot\cdot}}{rst(r-1)(t-1)}}.$$

10.29　10.9(a)節を参照しよう．層の大きさ，層の平均，層の分散が未知の，次のような2段標本抽出法を用いるとしよう．最初の標本は π_N からの大きさ r の単純無作為標本であり，それぞれ $\pi_{N_1},\cdots,\pi_{N_m}$ から $r_1,\cdots,r_m$ 個の要素を抽出する．ただし $r_1+\cdots+r_m=r$．次に $n_1=(r_1/r)n,\cdots,n_m=(r_m/r)n$ 個の要素がそれぞれ $\pi_{N_1},\cdots,\pi_{N_m}$ から無作為に取り出されるように，大きさ n の層化標本が π_N から抽出される．$\bar{x}_g$ を π_{N_g} からの大きさ n_g の標本平均とする．

$$\bar{x}'=\sum_g\frac{r_g}{r}\bar{x}_g$$

としよう．このとき，$\bar{x}'$ は π_N の平均 μ の不偏推定量となっていることを示せ．また $\bar{x}'$ の分散は

$$\sigma^2(\bar{x}')=\left[\frac{r(N-n-1)+n}{rnN}\right]\sigma_W^2+\frac{N-r}{rN}\sigma_B^2-\frac{n-r}{rnN(N-1)}\sum_{g=1}^{m}\sigma_g^2$$

となることを示せ．

10.30　(続き) N が $1/N$ 次の項を無視できるほど十分に大きいならば，

$$\sigma^2(\bar{x}') \cong \frac{\sigma^2_{W'}}{n} + \frac{\sigma^2_{B'}}{r}.$$

ただし $\sigma^2_{W'}$, $\sigma^2_{B'}$ は (10.10.14) で定義されている．

最初の標本では，1つの要素を抽出して観測する費用を c_1，2番目の標本では，1つの要素を抽出して観測する費用を c_2 としよう．この場合，総費用 $c = c_1 r + c_2 n$ が固定されたとき，$\sigma^2(\bar{x}')$ を最小にする r と n を選択せよ．

10.31 $(x_1, \cdots, x_n)$ が平均値 μ, 分散 σ^2 の分布からの標本で，u が μ に対する任意の線形不偏推定量ならば，$\bar{x}$ と u の間の相関係数は $\sigma(\bar{x})/\sigma(u)$ であることを示せ．

10.32 L_1 と L_2 を1つの標本から構成された母集団平均値の2つの線形不偏推定量とする．L_1, L_2 の分散を σ_1^2, σ_2^2 とし，L_1 と L_2 の間の相関係数を ρ とする．このとき，$c_1L_1 + c_2L_2$ が最小の分散を持つような定数 $c_1 > 0$, $c_2 > 0$, $c_1 + c_2 = 1$ を定めよ．

10.33 n, $\bar{x}$, s^2 を正規分布からの標本の大きさ，平均，分散とすれば，同じ正規分布から次に独立に抽出された x が区間 $\bar{x} \pm t_{n-1,\gamma}\sqrt{(n+1)/n}$ に落ちる確率は γ であることを示せ．ただし，$t_{n-1,\gamma}$ は (10.2.10) で定義している．

10.34 モーメントに対するシェパードの修正　$(x_1, \cdots, x_n)$ は p.d.f. $f(x)$ からの標本とする．h を $(x_1, \cdots, x_n)$ と独立な確率変数とし，矩形分布 $R(0, \delta)$ を持つとする．I_α を区間 $(h + \delta\alpha - \frac{1}{2}\delta, h + \delta\alpha + \frac{1}{2}\delta]$, $\alpha = \cdots, -2, -1, 0, 1, 2, \cdots$, n_α を I_α に落ちる標本要素の個数としよう．このとき，

$$M'_{r,\delta} = \frac{1}{n}\sum_{\alpha} n_\alpha (h + \alpha\delta)^r$$

は，標本要素が勝手にとった（そして“ランダムに位置をとった”）起点から最も近い長さ δ の区間単位上に分布しているとき，その r 次の標本モーメントと考えられる．$M'_{r,\delta}$ の特性関数 $\varphi(t)$ は

$$\varphi(t) = \frac{1}{\delta}\int_{-\frac{1}{2}\delta}^{+\frac{1}{2}\delta}\left[\sum_{\alpha} p_\alpha e^{it(h+\alpha\delta)^r/n}\right]^n dh$$

で与えられることを示せ．ただし $p_\alpha = \int_{I_\alpha} f(x)\,dx$ である．この特性関数を用いて

$$\mathscr{E}(M'_{1,\delta}) = \mu'_1, \quad \mathscr{E}(M'_{2,\delta}) = \mu'_2 + \delta^2/12, \quad \mathscr{E}(M'_{3,\delta}) = \mu'_3 + \delta^2\mu'_1/4,$$

さらに，一般に

$$\mathscr{E}(M'_{r,\delta}) = \frac{1}{\delta(r+1)}\left[\mathscr{E}\left(x + \frac{1}{2}\delta\right)^{r+1} - \mathscr{E}\left(x - \frac{1}{2}\delta\right)^{r+1}\right]$$

を証明せよ．ただし $\mu'_1, \mu'_2, \cdots$ は x のモーメントであり，また

$$\mathscr{E}\left(x \pm \frac{1}{2}\delta\right)^{r+1} = \int_{-\infty}^{+\infty}\left(x \pm \frac{1}{2}\delta\right)^{r+1} f(x)\,dx.$$

このように，$M'_{1,\delta}$, $M'_{2,\delta} - \delta^2/12$, $M'_{3,\delta} - M'_{1,\delta}\delta^2/4$, $\cdots$ は $\mu'_1, \mu'_2, \mu'_3, \cdots$ の不偏推定量となっている．[Sheppard (1898)].

10.35 $f(x_1, \cdots, x_k)$ を k 次元区間 $I_k = \{(x_1, \cdots, x_k) : 0 \leqslant x_i \leqslant a_i,\ i = 1, \cdots, k\}$ のすべての点 $(x_1, \cdots, x_k)$ に対して $0 < f(x_1, \cdots, x_k) < a_{k+1}$ となるような関数とする．また次の積分が存在するものとする．

$$J = \int_0^{a_1} \cdots \int_0^{a_k} f(x_1, \cdots, x_k)\,dx_1 \cdots dx_k.$$

$x_{1\xi}, \cdots, x_{k+1\xi}$ を矩形分布 $R\left(\frac{1}{2}a_1, a_1\right), \cdots, R\left(\frac{1}{2}a_{k+1}, a_{k+1}\right)$, $\xi = 1, \cdots, n$ からの独立な確率変数としよう．z_ξ は $x_{k+1\xi} \leqslant f(x_{1\xi}, \cdots, x_{k\xi})$ のとき，値1を持ち，$x_{k+1\xi} > f(x_{1\xi}, \cdots, x_{k\xi})$ のとき，値0を持つような確率変数とする ($\xi = 1, \cdots, n$)．また

$$\bar{z} = \frac{1}{n}\sum_{\xi=1}^{n} z_\xi$$

とする．このとき，$(a_1 \cdots a_{k+1})\bar{z}$ は J に対する不偏推定量で

$$\frac{(a_1 \cdots a_{k+1})^2}{4n}$$

を越えない分散を持つことを証明せよ．

第11章　ノンパラメトリック推定

11.1 序　　論

第10章では，有限および無限母集団の平均，分散，共分散，回帰係数などの母集団パラメータに対する1次ないしは2次推定量を求める問題を考えた．このような推定量の基礎になる線形推定の過程は，これまでに見てきたように，線形推定量に含まれる確率変数について弱い仮定，すなわち確率変数の平均と共分散行列の有限性，を必要とするだけの比較的簡単な手順である．

本章では，対象となる母集団の c.d.f. の分位および分位関数に対する比較的簡単な推定量のクラスを考えよう．この推定量は，標本が取り出される母集団の c.d.f. の関数形には依存しないことがわかる．この型の推定量を第12章で論じられる**パラメトリック推定量**と対比して**ノンパラメトリック推定量**と呼ぶ．パラメトリック推定量は，第12章で見るように，ある特定の関数形を持ち，そのパラメータに依存する c.d.f. に含まれる未知パラメータの値を推定する問題を取り扱うときによく出てくる．

ノンパラメトリック推定に含まれる基本的な確率変数は標本によって定まる順序統計量と，この順序統計量によって定まるカバー[*)] である．順序統計量とそのカバーに関する標本論は8.7節で論じてきたが，これらの結果のいくつかを次節で用いよう．

11.2 分位に対する信頼区間

(a) 小標本の場合

連続な c.d.f. $F(x)$ を持つ無限母集団がある．まず，$F(x)$ からの標本の順序統計量か

*) 区間変量ともいう．[訳注]

ら，与えられた分位 $\underline{x}_p$ に対する信頼区間を定める問題を考えよう．

さらに詳しく述べるために，$(x_1, \cdots, x_n)$ を $F(x)$ からの標本とし，$(x_{(1)}, \cdots, x_{(n)})$ を8.7節で定義したように，$F(x)$ からの標本順序統計量とする．$x_{(k_1)}$ と $x_{(k_1+k_2)}$ を標本の任意の2つの順序統計量とする．このとき，**8.7.3**から，カバーの和である確率変数 $u = F(x_{(k_1)})$，$v = F(x_{(k_1+k_2)})$ は順序つき2変数ディレクレ分布 $D^*(k_1, k_2;\ n - k_1 - k_2 + 1)$ を持つことがわかる．それゆえ，(u, v) の p.e. は

(11.2.1)
$$f(u, v)\,du\,dv = \begin{cases} \dfrac{\Gamma(n+1)}{\Gamma(k_1)\Gamma(k_2)\Gamma(n-k_1-k_2+1)} \\ \quad \cdot u^{k_1-1}(v-u)^{k_2-1}(1-v)^{n-k_1-k_2}\,du\,dv, & 0 < u < v < 1 \\ 0, & \text{その他}. \end{cases}$$

さて，不等式

(11.2.2)
$$F(x_{(k_1)}) < p < F(x_{(k_1+k_2)})$$

を考えよう．$F(x)$ は連続だから，(11.2.2) が満足されるための必要十分条件は

(11.2.3)
$$x_{(k_1)} < \underline{x}_p < x_{(k_1+k_2)}.$$

したがって，(11.2.3) が成り立つ確率は (11.2.2) が成り立つ確率に等しい．ゆえに

(11.2.4)
$$P(x_{(k_1)} < \underline{x}_p < x_{(k_1+k_2)}) = \int_p^1\int_0^p f(u, v)\,du\,dv.$$

しかし

(11.2.5)
$$\int_p^1\int_0^p f(u, v)\,du\,dv = \int_0^p\int_u^1 f(u, v)\,dv\,du - \int_0^p\int_0^v f(u, v)\,du\,dv.$$

(11.2.5) の右辺の最初と2番目の積分にそれぞれ変数変換

$$\begin{cases} u = r \\ v = 1 - s(1-r), \end{cases} \qquad \begin{cases} u = rs \\ v = s \end{cases}$$

を施せば

(11.2.6)
$$P(x_{(k_1)} < \underline{x}_p < x_{(k_1+k_2)}) = I_p(k_1, n - k_1 + 1) - I_p(k_1 + k_2, n - k_1 - k_2 + 1).$$

ただし $I_p(\nu_1, \nu_2)$ は不完全ベータ関数

(11.2.7)
$$I_p(\nu_1, \nu_2) = \frac{\Gamma(\nu_1 + \nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)}\int_0^p x^{\nu_1-1}(1-x)^{\nu_2-1}\,dx$$

である．

特に，区間 $(x_{(k_1)}, x_{(k_1+k_2)})$ が分位 $\underline{x}_p$ を含む確率は $F(x)$ には関係がないこと，そしてそれゆえ，この区間は (11.2.6) の右辺で与えられる信頼係数を持つ $\underline{x}_p$ に対する信

頼区間であることに注意しよう．まとめると，

11.2.1　$(x_1, \cdots, x_n)$ が連続な c.d.f. $F(x)$ から抽出した標本で，$x_{(k_1)}$，$x_{(k_1+k_2)}$ がその標本の第 k_1 番目と第 (k_1+k_2) 番目の標本順序統計量ならば，$(x_{(k_1)}, x_{(k_1+k_2)})$ は分位 $\underline{x}_p$ に対する信頼係数 $I_p(k_1, n-k_1+1) - I_p(k_1+k_2, n-k_1-k_2+1)$ を持つ信頼区間である．

k_1 と k_2 は（正）整数であるから，信頼係数 γ の値が (11.2.6) の型で与えられるならば，この γ に対して2つの順序統計量 $x_{(k_1)}$ と $x_{(k_1+k_2)}$ を用いれば，$\underline{x}_p$ の信頼区間を求めることができる．もちろん，γ の値がある範囲内にあるとき，信頼係数が少なくとも γ であるような信頼区間を求めることができる．さらに，ここで議論している2つの順序統計量は任意である．すなわち，(11.2.6) は任意の2つの順序統計量 $x_{(k_1)}$，$x_{(k_1+k_2)}$ に対して成り立っている．与えられた γ に対して，順序統計量を選ぶとき，通常，その順位ができるだけ互いに接近しているように選ぶ．すなわち，k_2 をできるかぎり小さくするように k_1，k_2 を選ぶのである．たとえば，メディアン $\underline{x}_{0.5}$ に対する信頼区間を定める場合には，

(11.2.8)
$$P(x_{(k)} < \underline{x}_{0.5} < x_{(n-k+1)}) \geqq \gamma$$

を満たす k の最大値を選べばよい．実際，Nair (1940) は $n = 6, 7, \cdots, 81$，$\gamma = 0.95, 0.99$ に対して，(11.2.8) が成り立つ k の値を作表している．確率 $P(x_{(k)} < \underline{x}_{0.5} < x_{(n-k+1)})$ の正確な値は $1 - 2I_{0.5}(n-k+1, k)$ である．この結果は Thompson (1936) と Savur (1937) によって個々に求められた．

(b)　大標本の場合

$(x_1, \cdots, x_n)$ を連続な c.d.f. $F(x)$ からの標本とし，$\underline{x}_p$ を p 分位とする．n_1 が $\underline{x}_p$ より小さい標本成分の数ならば，n_1 は2項分布 $Bi(n, p)$ を持つ確率変数である．**9.2.1 a** より，大きな n に対して，n_1 は漸近的に $N(np, np(1-p))$ に従って分布することがわかる．よって，与えられた信頼係数 γ に対して

(11.2.9)
$$\lim_{n\to\infty} P\left(-y_\gamma < \frac{n_1 - np}{\sqrt{np(1-p)}} < +y_\gamma\right) = \gamma$$

が得られる．ただし

(11.2.10)
$$\frac{1}{\sqrt{2\pi}} \int_{-y_\gamma}^{y_\gamma} e^{-\frac{1}{2}y^2} dy = \gamma.$$

このように n が大きいとき，p の近似 $100\gamma\%$ 信頼区間 $(\underline{p}_\gamma, \overline{p}_\gamma)$ は，固定された n_1，n，y_γ に対して，不等式 (11.2.9) を満足するような p のすべての値の集合によって与えられる．すなわち，$\underline{p}_\gamma$，$\overline{p}_\gamma$ は p に関する方程式

(11.2.11)
$$\frac{(n_1 - np)^2}{np(1-p)} = y_\gamma^2$$

の2つの解である．

よって，大きな n に対しては

(11.2.12)
$$P(\underline{p}_\gamma < p < \overline{p}_\gamma) \cong \gamma$$

となる．これは

(11.2.13)
$$P(x_{([n\underline{p}_\gamma])} < \underline{x}_p < x_{([n\overline{p}_\gamma])}) \cong \gamma$$

と同値である．ゆえに，2つの順序統計量 $(x_{([n\underline{p}_\gamma])}, x_{([n\overline{p}_\gamma])})$ は分位 $\underline{x}_p$ に対する近似 $100\gamma\%$ 信頼区間を構成している．ただし，$[n\underline{p}_\gamma]$ と $[n\overline{p}_\gamma]$ はそれぞれ $n\underline{p}_\gamma$，$n\overline{p}_\gamma$ に最も近い整数である．

11.3　分位区間に対する信頼区間

$p_1 < p_2$ を任意の2つの分位とし，$x_{(k_1)}$，$x_{(k_1+k_2)}$ を大きさ n の標本における順序統計量とする．このとき，(11.2.4) と同様にすれば

(11.3.1)
$$P(x_{(k_1)} < \underline{x}_{p_1} < \underline{x}_{p_2} < x_{(k_1+k_2)}) = \int_0^{p_1} \int_{p_2}^1 f(u, v)\, du\, dv$$

となる．この左辺は

(11.3.2)
$$P((\underline{x}_{p_1}, \underline{x}_{p_2}) \subset (x_{(k_1)}, x_{(k_1+k_2)})) = \frac{n!}{k_2!} \sum_{i=0}^{k_1-1} \frac{(-1)^i p_1^{i+k_1}}{i!(n-k_2-i)!} I_{p_1}(n-k_1-k_2+1; k_1-i)$$

と書き直せる．与えられた γ に対して，(11.3.2) で表わされる確率が値 γ を持つ k_1，k_2，n の値が存在すれば，$(x_{(k_1)}, x_{(k_1+k_2)})$ を分位区間 $(\underline{x}_{p_1}, \underline{x}_{p_2})$ に対する $100\gamma\%$ **外信頼区間**という．

同様に，$(\underline{x}_{p_1}, \underline{x}_{p_2})$ に対する $100\gamma\%$ **内信頼区間** $(x_{(k_1)}, x_{(k_1+k_2)})$ も次の関係式を考慮することによって得られる，

(11.3.3)
$$P(\underline{x}_{p_1} < x_{(k_1)} < x_{(k_1+k_2)} < \underline{x}_{p_2}) = \int_{p_1}^{p_2} \int_{p_1}^{v} f(u, v)\, du\, dv.$$

ただし，k_1，k_2，n は (11.3.3) で表わされる確率が γ に等しくなるような値を持ち，$f(u,v)$ は (11.2.1) で与えられている．

外（内）信頼区間の場合，k_1，k_2 の“最良の”選択は k_2 ができるかぎり小さい（大きい）ものと考えられる．たとえば，$p_1=p$，$p_2=1-p$，$2p<1$ のとき，内信頼区間および外信頼区間の両方に対する k_1 と k_2 のこの意味での“最良の”選択は，それぞれ $x_{(c)}$ と $x_{(n-c+1)}$ の形をとることが証明できる．

11.4　有限母集団における分位に対する信頼区間

π_N が有限母集団で，その要素が異なった x 値 $x_{o1}<\cdots<x_{oN}$ を持つとする．8.8節で，π_N からの大きさ n の標本における第 k 番目の順序統計量 $x_{(k)}$ の p.f. は次のようになることを示した．

$$p_{N,n,k}(t)=\frac{\binom{t-1}{k-1}\binom{N-t}{n-k}}{\binom{N}{n}}. \tag{11.4.1}$$

このとき，$x_{(k)}$ の質点は x_{ot}，$t=k,k+1,\cdots,N-n+k$ である．さて，ここで任意に固定された t の値 t' に対して

$$P(x_{(k)}\leqslant x_{ot'})=\sum_{t=k}^{t'}p_{N,n,k}(t) \tag{11.4.2}$$

である．固定された N，n，t' および $\gamma>0$ に対して

$$\sum_{t=k'}^{t'}p_{N,n,k'}(t)\geqslant\gamma \tag{11.4.3}$$

となる最大な k（それを k' とする）があるとき，$x_{(k')}$ は $x_{ot'}$ に対する“最良の”下位 $100\gamma\%$ 信頼限界とみなされる．$x_{ot'}$ は母集団 π_N の (t'/N) 分位と考えてよい．N，n，t'，$1-\gamma$ の値が問題にならないほど小さい場合以外は，このような下位信頼限界が存在することが示される．

同様に

$$\sum_{t=t'}^{N-n+k''}p_{N,n,k''}(t)\geqslant\gamma \tag{11.4.4}$$

となる最小の k（k'' とする）を選べば，$x_{ot'}$ の“最良の”上位 $100\gamma\%$ 信頼区間が得られる．

$x_{ot'}$ に対する“最良の” $100\gamma\%$ 信頼区間，すなわち，上位信頼限界と下位信頼限界の両方も求めることができる．しかし，ここに含まれる p.f. は p.f. $P_{N,n,k}(t)$ よりも繁雑なので書かないでおこう．しかしながら，母集団 π_N が相異なる x 値 $x_{o1}<\cdots<x_{oN}$ を持ち，$x_{(k_1)}$，$x_{(k_1+k_2)}$ が π_N からの大きさ n の標本の順序統計量ならば，

$$\lim_{N\to\infty}P(x_{(k_1)}<x_{o(Np)}<x_{(k_1+k_2)})=I_p(k_1,n-k_1+1)-I_p(k_1+k_2,n-k_1-k_2+1) \tag{11.4.5}$$

が示される．したがって，N が大きいとき，$(x_{(k_1)},x_{(k_1+k_2)})$ は (11.4.5) の右辺で近似される係数を持つ $x_{o(Np)}$ に対する信頼区間である．

11.5　許　容　限　界

(a)　小標本の場合

$(x_1,\cdots,x_n)$ を連続な c.d.f. $F(x)$ からの標本とする．いま，$L_1(x_1,\cdots,x_n)<L_2(x_1,\cdots,x_n)$ が $(x_1,\cdots,x_n)$ に関する任意の2つの観測可能な対称式で，$F(L_2)-F(L_1)$ の分布が $F(x)$ に関係なく，$0<\beta<1$ に対して

$$P[(F(L_2)-F(L_1))\geqslant\beta]=\gamma \tag{11.5.1}$$

であるとする．このとき，L_1 と L_2 を確率水準 γ での分布によらない $100\beta\%$ 許容限界と呼ぶ．この概念は Shewhart (1931) によるものである．

Wilks (1941, 1942) が指摘したように，順序統計量は $F(x)$ が連続ならばいつでも分布によらない許容限界として使用できる．その理由は，まず $(x_{(1)},\cdots,x_{(n)})$ を未知の連続な c.d.f. $F(x)$ から抽出した標本順序統計量とする．このとき，任意の2つの順序統計量 $x_{(k_1)}$，$x_{(k_1+k_2)}$ に対して，区間 $(x_{(k_1)},x_{(k_1+k_2)})$ に含まれる母集団分布による確率 U_{k_2} は1つの確率変数である．8.7(b)節より，U_{k_2} は k_2 個のカバーの和になり，

$$U_{k_2}=F(x_{(k_1+k_2)})-F(x_{(k_1)}) \tag{11.5.2}$$

である．

8.7.6 から，U_{k_2} はベータ分布 $Be(k_2,n-k_2+1)$ を持ち，母集団 c.d.f. $F(x)$ には依存しないことがわかる．さてここで，与えられた $\beta>0$，$\gamma>0$ の値に対して

(11.5.3)
$$P(U_{k_2} \geqq \beta) = \gamma$$
となる n, k_1, k_2 が存在すると仮定しよう．このとき，$x_{(k_1)}$ と $x_{(k_1+k_2)}$ は確率水準 γ での分布によらない $100\beta\%$ 許容限界となっている．U_{k_2} は (11.5.1) における $F(L_2)-F(L_1)$ の特別な形である．Robbins (1944 b) は，$F(x)$ が絶対連続，すなわち導関数を持つならば，L_1, L_2 は必然的に順序統計量となることを示した．

$k_1=c$, $k_2=n-2c+1$ を仮定して，順序統計量から対称に選んで区間 $(x_{(c)}, x_{(n-2c+1)})$ を構成しよう．ここで不完全ベータ関数の記号を用いれば，(11.5.3) は

(11.5.4)
$$1-I_\beta(n-2c+1, 2c)=\gamma$$
となる．c を固定すると，この等式が厳密に成り立つような，標本の大きさ n は存在しないかも知れない．しかしながら，任意の $\varepsilon>0$ に対してベータ分布 $Be(n-2c+1, 2c)$ から計算される区間 $(1-\varepsilon, 1)$ 上の確率は n を十分大きくとれば，任意に1に近づけることができるから，式

(11.5.5)
$$1-I_\beta(n-2c+1, 2c) \geqq \gamma$$
を満たす最小の n は存在する．たとえば，$\beta=0.99$, $\gamma=0.95$, $c=1$ (すなわち，$k_1=1$, $k_2=n-1$) とすれば，$n=473$ となる．Murphy (1948) は $n=1\,(1)\,10\,(10)\,100\,(100)\,500$, $\gamma=0.90, 0.95, 0.99$, $n-k_2+1=m=1\,(1)\,6\,(2)\,10\,(5)\,30\,(10)\,60\,(20)\,100$ の場合に (11.5.3) が成り立つ β の値を作表している．Somerville (1958) は Murphy (1948) の結果を数表に拡張している．

連続な c.d.f. を持つ1次元分布に対する分布によらない許容限界の概念は，連続な c.d.f. を持つ多次元分布の場合に拡張できる．いま，連続な k 次元 c.d.f. $F(x_1, \cdots, x_k)$ を持つ無限母集団から大きさ n の標本をとったとしよう．$(x_1, \cdots, x_k)$ の k 次元標本空間が 8.7(c) 節で述べた順序関数によって，$n+1$ 個の互いに排反でしかも全空間を尽す標本ブロック $B_k^{(1)}, \cdots, B_k^{(n+1)}$ に切断されたとする．これらの標本ブロックから r 個を選び出す任意の規則を考えよう．このブロックの和を T_r とする．選ばれたブロックに対するカバーの和を U_r とする．すなわち

(11.5.6)
$$U_r=\int_{T_r} dF(x_1, \cdots, x_k).$$
このとき，U_r は1つの確率変数となり，**8.7.9** より U_r はベータ分布 $Be(r, n-r+1)$ を持つ．さて，与えられた $\beta>0$, $\gamma>0$ に対して

(11.5.7)
$$P(U_r \geqq \beta)=\gamma$$

ならば，U_r は確率水準 γ での 分布によらない $100\beta\%$ 許容域となる．任意に固定した正整数 c に対して，$r=n-2c+1$ と選べば (11.5.7) は (11.5.4) になり，1次元許容限界に対する同じ式を得る．

(b)　有限母集団の場合

t' を $t'/N=(1-\beta)$ と選べば，(11.4.2) は

(11.5.8)
$$P(x_{(k')} \leqq x_{o\,(N(1-\beta))}) \geqq \gamma.$$
この場合，$x_{(k')}$ は確率水準 γ での $100\beta\%$ 下位許容限界とみなされる．すなわち，π_N において $x_{(k')}$ を越える x 値を持つ要素の割合が β となる確率は少なくとも γ である．

同様に，$t'=\beta N$ のとき，(11.4.4) は

(11.5.9)
$$P(x_{(k'')} \geqq x_{o\,(N\beta)}) \geqq \gamma$$
となり，$x_{(k'')}$ は確率水準 γ での $100\beta\%$ 上位許容限界となる．

確率水準 γ での $100\beta\%$ 許容区間をどのように定義すればよいかは明らかである．

異なる x 値を持つ大母集団 π_N については，π_N において x 値が区間 $(x_{(k_1)}, x_{(k_1+k_2)})$ にはいる割合を $U_{k_2,N}$ とすれば，$N \longrightarrow \infty$ のとき，$U_{k_2,N}$ の極限分布は $Be(k_2, n-k_2+1)$ となることが示される．したがって，大きな N に対して，この区間は確率水準 γ での近似 $100\beta\%$ 許容区間である．

11.6　連続な分布関数に対する片側信頼水準

順序統計量に関する標本論を統計的推定論に適用する場合，基本的な問題の1つに，連続な c.d.f. $F(x)$ を持つ母集団からの標本順序統計量から，$F_n^+(x)$, $F_n^-(x)$ を以下の各条件を満たすように構成する問題がある．すなわち，与えられた γ に対して

(11.6.1)
$$P(F_n^+(x) \geqq F(x);\ \text{すべての } x \text{ に対して})=\gamma$$
$$P(F_n^-(x) \leqq F(x);\ \text{すべての } x \text{ に対して})=\gamma.$$
このような関数 $F_n^+(x)$, $F_n^-(x)$ をそれぞれ，$F(x)$ に対する上位，下位 $100\gamma\%$ 信頼水準と呼ぶ．

大きな n に対するこの問題の漸近解は，Smirnov (1939 a) によって最初に得られた．その後，Smirnov (1944), Birnbaum と Tingy (1951) 等も，ここで述べようとしてい

る確率 (11.6.1) に対する簡単な表現を見い出している．

$(x_{(1)}, \cdots, x_{(n)})$ を，c.d.f. $F(x)$ を持つ無限母集団からの標本順序統計量としよう．経験 c.d.f. $F_n(x)$ はこの順序統計量から次のように構成される．

(11.6.2)
$$F_n(x) = \begin{cases} 0, & x < x_{(1)} \\ \dfrac{\xi - 1}{n}, & x_{(\xi-1)} \leqslant x < x_{(\xi)}, \quad \xi = 2, \cdots, n \\ 1, & x \geqslant x_{(n)}. \end{cases}$$

$0 < d \leqslant 1$ および任意の x の値に対して

(11.6.3)
$$F_n^+(x, d) = \min\,[F_n(x) + d\,;\, 1]$$
$$F_n^-(x, d) = \max\,[F_n(x) - d\,;\, 0]$$

と置こう．そして

(11.6.4)
$$D_n^+(d) = \inf_x\,(F_n^+(x, d) - F(x))$$
$$D_n^-(d) = \sup_x\,(F_n^-(x, d) - F(x))$$

とする．

$P(D_n^+(d) > 0)$ は $F_n(x) + d$ のグラフが $F(x)$ のグラフと交わることがない確率である．同様に $P(D_n^-(d) < 0)$ は $F_n(x) - d$ のグラフが $F(x)$ のグラフと交わらない確率である．

主な結果とまとめると，

11.6.1　$(x_{(1)}, \cdots, x_{(n)})$ を連続な c.d.f. $F(x)$ からの標本順序統計量とし，$F_n^+(x, d)$, $F_n^-(x, d)$ を (11.6.3) のように，経験 c.d.f. $F_n(x)$ から構成する．このとき，$P(D_n^+(d) > 0) = P(D_n^-(d) < 0) = P_n(d)$ が成り立つ．ただし，

(11.6.5)
$$P_n(d) = 1 - d \sum_{i=1}^{[n(1-d)]} \binom{n}{i} \left(1 - d - \frac{i}{n}\right)^{n-i} \left(d + \frac{i}{n}\right)^{i-1}$$

で，$[n(1-d)]$ は $n(1-d)$ 以下の最大整数である．

11.6.1 の証明は，まず次で定義する新しい確率変数 $y_{(1)}, \cdots, y_{(n)}$ に変換することから始めよう．

(11.6.6)
$$y_{(\xi)} = F(x_{(\xi)}), \qquad \xi = 1, \cdots, n.$$

このとき，(8.7.2) から $y_{(1)}, \cdots, y_{(n)}$ は，領域 $0 < y_{(1)} < \cdots < y_{(n)} < 1$ の内部では p.e.

(11.6.7)
$$n!\, dy_{(1)} \cdots dy_{(n)}$$

を持ち，その他では0となる．

まず $P(D_n^+(d) > 0)$ を考えよう．順序統計量 $x_{(1)}, \cdots, x_{(n)}$ が不等式 $D_n^+(d) > 0$ を満たすための必要十分条件は順序統計量 $y_{(1)}, \cdots, y_{(n)}$ が不等式

(11.6.8)
$$y_{(\xi-1)} < y_{(\xi)} < \frac{\xi - 1}{n} + d, \qquad \xi = 1, \cdots, k+1$$
$$y_{(\xi-1)} < y_{(\xi)} < 1, \qquad \xi = k+2, \cdots, n$$

を満たすことである．ただし，k は $n(1-d)$ 以下の最大整数であり，$y_{(0)} = 0$ である．同様に，$D_n^-(d) < 0$ が満たされるための必要十分条件は

(11.6.9)
$$\frac{\xi}{n} - d < y_{(\xi)} < y_{(\xi+1)}, \qquad \xi = n-k, \cdots, n$$
$$0 < y_{(\xi)} < y_{(\xi+1)}, \qquad \xi = 1, \cdots, n-k-1.$$

ただし，$y_{(n+1)} = 1$．しかし，ここで変数変換 $y_{(\xi)} = 1 - y'_{(n-\xi+1)}$, $\xi = 1, \cdots, n$ を行なえば，順序統計量 $y'_{(1)}, \cdots, y'_{(n)}$ は p.e. (11.6.7) を持ち，不等式 (11.6.9) は (11.6.8) になる．したがって，$P(D_n^+(d) > 0) = P(D_n^-(d) < 0)$．ゆえに，この確率の共通な値（これを $P_n(d)$ とすれば）は，p.e. (11.6.7) を不等式 (11.6.8) で定まる $0 < y_{(1)} < \cdots < y_{(n)} < 1$ 内の領域上で積分することによって得られる．すなわち

(11.6.10)
$$P_n(d) = n!\, G_n(k, d)$$

ただし

(11.6.11)
$$G_n(k, d) = \int_0^d \int_{y_{(1)}}^{1/n+d} \cdots \int_{y_{(k)}}^{k/n+d} \int_{y_{(k+1)}}^1 \cdots \int_{y_{(n-1)}}^1 dy_{(n)} \cdots dy_{(1)}.$$

$y_{(n)}, \cdots, y_{(k+2)}$ に関して積分すれば，

(11.6.12)
$$G_n(k, d) = \int_0^d \int_{y_{(1)}}^{1/n+d} \cdots \int_{y_{(k)}}^{k/n+d} \frac{(1 - y_{(k+1)})^{n-k-1}}{(n-k-1)!}\, dy_{(k+1)} \cdots dy_{(1)}.$$

ここからは，$G_n(k, d)$ の値を帰納法で評価する．すなわち，$G_n(k+1, d)$ において $y_{(k+2)}$ に関して積分すれば

(11.6.13)
$$G_n(k+1, d) = G_n(k, d) - \frac{\left(1 - d - \dfrac{k+1}{n}\right)^{n-k-1}}{(n-k-1)!} H_n(k, d).$$

ただし

(11.6.14)
$$H_n(k, d) = \int_0^d \int_{y_{(1)}}^{1/n+d} \cdots \int_{y_{(k)}}^{k/n+d} dy_{(k+1)} \cdots dy_{(1)}.$$

数学的帰納法の詳細は省略するが

(11.6.15)　$$H_n(k,d) = -\frac{d}{(k+1)!}\sum_{i=1}^{k+1}(-1)^i\binom{k+1}{i}\left(d+\frac{k+1-i}{n}\right)^k$$

が得られる．これは次のように書ける．

(11.6.16)　$$H_n(k,d) = -\frac{d}{(k+1)!}\frac{\partial^k}{\partial t^k}\{e^{(d+(k+1)/n)t}[(1-e^{-t/n})^{k+1}-1]\}_{t=0}.$$

しかし

(11.6.17)　$$\left\{\frac{\partial^k}{\partial t^k}[e^{(d/(k+1)+1/n)t}-e^{(d/(k+1))t}]^{k+1}\right\}_{t=0}=0$$

に注意すれば

(11.6.18)　$$H_n(k,d) = \frac{d}{(k+1)!}\left\{\frac{\partial^k}{\partial t^k}e^{(d+(k+1)/n)t}\right\}_{t=0} = \frac{d}{(k+1)!}\left(d+\frac{k+1}{n}\right)^k.$$

$H_n(k,d)$ に対するこの値を (11.6.13) に代入して，

(11.6.19)　$$G_n(0,d) = \frac{1-(1-d)^n}{n!}$$

を用いれば，帰納法によって $G_n(k,d)$ の値が求められる．これを (11.6.10) に代入し，k が $n(1-d)$ 以下の最大整数であることに注目すれば，$P_n(d)$ に対する (11.6.5) の式が出る．これで **11.6.1** の証明は終わる．

Birnbaum と Tingy (1951) は $\gamma = 0.90, 0.95, 0.99, 0.999$,　$n = 5, 8, 10, 20, 40, 50$ に対して $P_n(d) = \gamma$ なる d の値を表にしている．50より大きい n の値についてはスミルノフの漸近的近似 (11.6.22) が d の値に対するよい近似値を与えている．

さて，ここで $n \longrightarrow \infty$ として，$P_n(d)$ の極限を考えてみよう．$d = \lambda/\sqrt{n}$ とおけば，(11.6.5) は

(11.6.20)

$$P_n\left(\frac{\lambda}{\sqrt{n}}\right) = 1-\lambda\left\{\sum_{i=1}^{[n-\sqrt{n}\lambda]}\sqrt{n}\binom{n}{i}\left(1-\frac{\lambda}{\sqrt{n}}-\frac{i}{n}\right)^{n-i}\left(\frac{\lambda}{\sqrt{n}}+\frac{i}{n}\right)^{i-1}\Delta y\right\}$$

になる．ただし，$\Delta y = 1/n$ である．大きな階乗に対するスターリングの近似式(7.6.29)を用いれば，{　} 内の和は $n \longrightarrow \infty$ のとき積分

(11.6.21)　$$\frac{1}{\sqrt{2\pi}}\int_0^1\frac{1}{\sqrt{y(1-y)}}e^{-\lambda^2/2y(1-y)}\,dy$$

に収束する．これを積分すれば，

$$\frac{1}{\lambda}e^{-2\lambda^2}.$$

したがって，次の結果が得られる．

11.6.2　**11.6.1** の条件の下では，

(11.6.22)　$$\lim_{n\to\infty}P_n\left(\frac{\lambda}{\sqrt{n}}\right) = 1-e^{-2\lambda^2}.$$

この結果は最初 Smirnov (1939 *a*) が得た．しかし，(11.6.5) から直接に導き出したのは Dempster (1955) である．

11.7　連続な分布関数に対する信頼帯

$F(x)$ に対する両側信頼水準，すなわち，信頼帯を定める問題，いい換えれば，任意の d に対して，確率

(11.7.1)　$$P(D_n < d)$$

の値を決める問題は，前節で述べた片側信頼水準の問題よりも難しい．ここに

$$D_n = \sup_x |F_n(x) - F(x)|$$

である．Kolmogorov (1933 *b*) は大きな n に対するこの問題の漸近的な解を与え，有限な n に対して，(11.7.1) の値が計算できる再起公式を導いている．Wald と Wolfowitz (1939) は有限な n に対する解を行列式の形で与えている．最近では，Massey (1950) が，n が有限で d が $1/n$ の倍数のとき，この問題に対するかなり簡単な解を求めている．彼の解は再起公式の形をとっているが，ここではそれを考えてみよう．マーセイの結果は次のように述べることができる．

11.7.1　$(x_{(1)}, \cdots, x_{(n)})$ を連続な c.d.f. $F(x)$ からの大きさ n の標本順序統計量とする．また

(11.7.2)　$$D_n = \sup_x |F_n(x) - F(x)|$$

とする．ただし $F_n(x)$ は (11.6.2) で定義されている．このとき

(11.7.3)　$$P\left(D_n < \frac{k}{n}\right) = \frac{n!}{n^n}U(k,n),\qquad k = 1, \cdots, n-1$$

である．ただし，$U(j, m+1)$, $j = 1, \cdots, 2k-1$, $m = 0, 1, \cdots, n-1$ は，等式系

(11.7.4)
$$U(j, m+1) = \sum_{i=1}^{j+1} \frac{U(i, m)}{(j+1-i)!}$$

を境界条件

(11.7.5)
$$\begin{aligned} &U(i, m) = 0, && i \geqslant m+k \\ &U(i, 0) = 0, && i = 1, \cdots, k-1 \\ &U(k, 0) = 1 \end{aligned}$$

の下で満足するものとする.

この証明にはまず，$(x_1, \cdots, x_n)$ を連続な c.d.f. $F(x)$ からの標本とし，便宜上，確率変数 $F(x_\xi)$ を y_ξ, $\xi = 1, \cdots, n$ で表わすことにする．このとき，y_ξ は区間 $(0,1)$ 上の矩形分布，すなわち，$R\left(\frac{1}{2}, 1\right)$ を持つ．$G(y)$ を y の c.d.f. とし，$G_n(y)$ を順序統計量 $(y_{(1)}, \cdots, y_{(n)})$ から構成された経験 c.d.f. とする．このとき，もちろん

(11.7.6)
$$P\left(D_n < \frac{k}{n}\right) = P\left(\sup_y |G_n(y) - G(y)| < \frac{k}{n}\right)$$

である．さて，区間 $(0, 1]$ を長さが等しい n 個の区間 $I_\xi = ((\xi-1)/n, \xi/n]$, $\xi = 1, \cdots, n$ に分割しよう．$(r_1, \cdots, r_n)$ を標本 $(y_1, \cdots, y_n)$ の要素のうち，それぞれ $I_1, \cdots, I_n$ に落ちる個数を表わす $(r_1 + \cdots + r_n = n$ で退化する) 確率変数としよう．もちろん，r は $(n-1)$ 次元多項分布 $M(n; 1/n, \cdots, 1/n)$ を持ち，その p.f. は

(11.7.7)
$$p(r_1, \cdots, r_n) = \frac{n!}{r_1! \cdots r_n!\, n^n}$$

である．いま，$(r_1, \cdots, r_n)$ は $G_n(y)$ を一意的に決めている．したがって，$P(D_n < k/n)$ の値はすべての y に対して $|G_n(y) - G(y)| < k/n$ を満たす $(r_1, \cdots, r_n)$ の標本空間内の点全体にわたる $p(r_1, \cdots, r_n)$ の和によって定まる．$G_n(y)$ のグラフ（これは，$y = 1/n, 2/n, \cdots, (m+1)/n$ に対して $|G_n(y) - G(y)| < k/n$ となる帯 E の中に完全にはいっている）を左から右へ追って行けば，このグラフは必ず点 $(m/n, (m-k+i)/n)$, $i = 1, \cdots, 2k-1$ のうち1点を通るはずである．これらの点をそれぞれ $A(i, m)$, $i = 1, \cdots, 2k-1$ としよう．さて，

(11.7.8)
$$U(i, m) = \sum\nolimits_{(i)} \frac{1}{r_1! \cdots r_m!}$$

としよう．ただし，$\sum_{(i)}$ は，$G_n(y)$ のグラフが帯 E の中にとどまっている間に $A(i, m)$ に達するときの $(r_1, \cdots, r_m)$ の値の組全体にわたる和を表わしている．すると，$G_n(y)$ は非減少だから，そのパスが $A(j, m+1)$ に到達できるのは，点 $A(1, m), A(2, m), \cdots, A(j+1, m)$ の1点を通り，r_{m+1} がそれぞれ値 $j, j-1, \cdots, 1, 0$ をとったときに限る．したがって，$j = 1, 2, \cdots, 2k-1$ および $m = 0, 1, \cdots, n-1$ に対して

(11.7.9)
$$U(j, m+1) = \frac{U(1, m)}{j!} + \frac{U(2, m)}{(j-1)!} + \cdots + \frac{U(j+1, m)}{0!}.$$

ただし，$i \geqslant m+k$ ならば $U(i, m) = 0$ である．さらに $U(i, m)$ が次の境界条件を持つことは明らかである．

(11.7.10)
$$\begin{aligned} &U(i, 0) = 0, && i = 1, \cdots, k-1 \\ &U(k, 0) = 1. \end{aligned}$$

$U(k, n)$ について差分方程式系 (11.7.9) を条件 (11.7.8) の下で解けば，公式 (11.7.3) によって，$P(D_n < k/n)$ が得られる．これで，**11.7.1** の議論が完了する．

マーセイは $n = 5(5)80$ および $k = 1(1)9$ に対する $P(D_n < k/n)$ の値を作表している．彼はこれらの値から $n = 10(10)80$ および $\lambda = 0.9(0.1)1.40$ に対する $P(D_n < \lambda/\sqrt{n})$ の値を補間法で計算している．

前に述べた Kolmogorov (1933 *b*) の漸近的結果は，**11.7.1** の条件下で

(11.7.11)
$$\lim_{n\to\infty} P\left(D_n < \frac{\lambda}{\sqrt{n}}\right) = \sum_{i=-\infty}^{+\infty} (-1)^i e^{-2i^2\lambda^2}$$

となる．この結果のコルモゴロフによる証明は複雑である．簡単な証明は Feller (1948) が行なっており，Doob (1949) と Donsker (1952) は正規確率過程論を用いた証明を与えている．Darling (1957) はこの問題や関連する問題を含むような包括的な論文を発表している．それには広範囲にわたる文献が載っている．

特に (11.7.11) より，任意の $\varepsilon > 0$ に対して

(11.7.12)
$$\lim_{n\to\infty} P(D_n < \varepsilon) = 1$$

がわかる．すなわち，$F_n(x)$ は，$n \longrightarrow \infty$ のとき x に関して一様に $F(x)$ に確率収束している．

問　題

11.1　p.e. が (11.2.1) であるような (u, v) の標本空間において，E が $u < p$ となる点集合で，F が $v > p$ となる集合であるとき，基本的な確率法則 $P(E \cup F) = P(E) + P(F) - P(E \cap F)$ から (11.2.6) を導き出せ．

11.2　(11.2.8) において，次を証明せよ．

$$P(x_{(k)}<\underline{x}_{0.5}<x_{(n-k+1)})=1-2I_{\frac{1}{2}}(n-k+1,k).$$

11.3　$(x_{(1)},\cdots,x_{(n)})$ を連続な c.d.f. $F(x)$ からの標本順序統計量とするとき

$$F(x_{(n)})-F(x_{(1)})>\beta$$

となる確率は $1-n\beta^{n-1}+(n-1)\beta^n$ であることを示せ．(すなわち，その標本範囲に含まれる母集団の割合は，確率 $1-n\beta^{n-1}+(n-1)\beta^n$ を持って β を越える)．

11.4 (続き)　$\delta_1+\delta_2<1$ なる任意の正の値 δ_1, δ_2 に対し，次の不等式の両方が成り立つ確率は $1-(1-\delta_1)^n-(1-\delta_2)^n+(1-\delta_1-\delta_2)^n$ であることを示せ．

$$F(x_{(1)})<\delta_1,\qquad 1-F(x_{(n)})<\delta_2.$$

11.5 (続き)　任意に固定された整数 $k<n/2$ に対して，r があらゆる可能な値 (正，負の整数または0) をとるとき，$P(x_{(k+r)}<\underline{x}_{0.5}<x_{(n-k+r+1)})$ が最大値をとるのは $r=0$ のときであり，このとき，$F(\underline{x}_{0.5})=0.5$ であることを示せ．

11.6 (続き)　次を証明せよ．

$$P(x_{(k)}<\underline{x}_p)=I_p(k,n-k+1)$$

ただし $F(\underline{x}_p)=p$.

11.7 (続き)　$(x_{(k)},x_{(n-k+1)})$ は同じ分布からの大きさ $2m+1$ の独立な標本のメディアンに対する信頼係数

$$\gamma_m=\sum_{t=k}^{n-k}p_m(t)$$

を持つ信頼区間と考えてもよいことを示せ．ただし

$$p_m(t)=\frac{(m+1)}{(m+t+1)}\cdot\frac{\binom{2m+1}{m}\binom{n}{t}}{\binom{2m+n+1}{m+t+1}}.$$

$m\longrightarrow\infty$ のとき，γ_m の極限は $1-2I_{0.5}(n-k+1,k)$ であることを示せ．

11.8　(11.2.6) において，$p=0.5$ ならば

$$P(x_{(k_1)}<\underline{x}_{0.5}<x_{(k_1+k_2)})=\left(\frac{1}{2}\right)^n\sum_{\xi=k_1}^{k_1+k_2-1}\binom{n}{\xi}$$

となることを示せ．

11.9　$(x_{(1)},\cdots,x_{(n)})$ を要素が相異なる x 値を持つ大きさ $2m+1$ の有限母集団からの大きさ n の標本順序統計量とする．このとき，$(x_{(k)},x_{(n-k+1)})$ は前述の問題で与えたのとまったく同じ信頼係数 γ_m を持つ母集団メディアンに対する信頼区間であることを示せ．

11.10　大きさ n の標本を連続な c.d.f. から抽出したとき，さらに大きさ m の独立な標本の中の t 個の x が最初の標本範囲内に落ちる確率は

$$\frac{n(n-1)}{(m+n)}\binom{m}{t}\Big/\binom{m+n-1}{m-t+1}$$

であることを証明せよ．

11.11 (続き)　$U_{1,n,m}$ を2番目の標本の x が最初の標本範囲内にはいる割合とする．$m\longrightarrow\infty$ のときの $U_{1,n,m}$ の極限分布はベータ分布 $Be(n-1,2)$ であることを示せ．一般に，$x_{(k_1)}$, $x_{(k_1+k_2)}$ を最初の標本の順序統計量とし，$U_{k_1,k_1+k_2,m}$ を2番目の標本の x で，区間 $(x_{(k_1)},x_{(k_1+k_2)})$ 内にあるものの割合とすれば，$U_{k_1,k_1+k_2,m}$ の極限分布はベータ分布 $Be(k_2,n-k_2+1)$ であることを示せ．

11.12　連続な2次元 c.d.f. からの大きさ n の標本が xy 平面の n 個の点で表わされているとする．それぞれ最小，最大の x 座標を持つ点を通る垂線を引く．次に残りの $n-2$ 個の点のうちで，最小，最大の y 座標を持つ点を通る水平線を引く．これらの4直線で囲まれる矩形内に含まれる母集団の割合 U を考えよう．(U はこの矩形のカバーになっている)．カバー U が β を越える確率は

$$1-4\binom{n}{4}\beta^{n-3}\left[\frac{1}{n-3}-3\frac{\beta}{n-2}+3\frac{\beta^2}{n-1}-\frac{\beta^3}{n}\right]$$

であることを示せ．

11.13　U_{n+1-k} の定義について 11.5(a) 節を参照して，k を固定して $n(1-\beta)\longrightarrow\lambda$ となるように $n\longrightarrow\infty$, $\beta\longrightarrow1$ とすれば

$$\lim P(U_{n+1-k}>\beta)=1-e^{-\lambda}\left[1+\lambda+\frac{\lambda^2}{2!}+\cdots+\frac{\lambda^k}{k!}\right]$$

であることを示せ．

11.14　(11.4.5) を確かめよ．

11.15　$H_n(k,d)$ は (11.6.18) で与えられることを確かめよ．

11.16　(11.6.20) における｛　｝の量は $n\longrightarrow\infty$ のとき，積分 (11.6.21) に収束することを示せ．

11.17　$\bar{x}$, s^2 を $N(\mu,\sigma^2)$ からの大きさ n の標本の標本平均，標本分散とする．また $G(x)$ を $N(\mu,\sigma^2)$ の c.d.f. とする．(L_1,L_2) を区間 $\bar{x}\pm\lambda s$, $\lambda>0$ としよう．このとき，カバーとして，$G(L_2)-G(L_1)$ は λ のみに依存する分布を持つことを示せ．また $t_{n-1,\gamma}$ を (10.2.10) で定義するとき，$\lambda=t_{n-1,\gamma}\sqrt{(n+1)/n}$ に対して $\mathcal{E}(G(L_2)-G(L_1))=\gamma$ であることを示せ．[Wilks (1941), Wald と Wolfowitz (1946) を見よ]．

第12章　パラメトリック推定

統計的推定に関する多くの問題は特定の関数形 $F(x;\theta)$ の c.d.f. からの標本問題を取り扱っている．ここで，θ は未知の（実）パラメータで，その真値 θ_0 は $F(x;\theta)$ から抽出される標本 $(x_1,\cdots,x_n)$ の要素から推定される．真値 θ_0 は θ の値からなるパラメータ空間 Ω の点であり，Ω はユークリッド空間の開区間または領域（またはすべて）である．パラメータ空間 Ω は θ の**許容値の集合**とも呼ばれる．いうまでもなく，x と θ の両方とも多次元であってもよい．

パラメトリック推定量には 2 つの重要な型，すなわち**点推定量**と**区間推定量**がある．点推定量では，$F(x;\theta)$ におけるパラメータ θ の真値 θ_0 に対する推定量は観測可能な確率変数である．それを $\tilde{\theta}(x_1,\cdots,x_n)$ とする．これは標本要素 $(x_1,\cdots,x_n)$ の関数であり，この分布はある意味で θ の真値 θ_0 のまわりに集中している．線形推定において，点推定量の分散は集中度の尺度として合理的な基準になることがわかる．

区間推定の場合は，観測可能な 2 つの確率変数 $\underline{\theta}(x_1,\cdots,x_n)$，$\bar{\theta}(x_1,\cdots,x_n)$ に分けられる．これを $(\underline{\theta},\bar{\theta})$ と略して書く．ただし $\underline{\theta}<\bar{\theta}$．また確率区間 $(\underline{\theta},\bar{\theta})$ が θ_0 を含む確率が与えられ，ある意味でできるだけ短くしてある．もちろんこの概念は $F(x;\theta)$ の x, θ（または両方）が多次元の場合にも拡張できる．“パラメータ θ の真値 θ_0 に対する推定量 $\tilde{\theta}$” というかわりに，“θ に対する推定量 $\tilde{\theta}$” といおう．

第 10 章ですでに考察した線形推定量は点推定量の 1 例である．10.2 (c), 10.4, 10.5, 10.6, 11.2 (a) で区間推定量の例を見てきた．そこでは標本は正規分布から抽出されていた．

本章では，点推定量および区間推定量の概念を無限母集団からの標本に対するパラメトリック推定の一般的な枠組の中で議論し，大標本に対する漸近的な結果のみならず有限標本に関する基本的な結果もいくつか示そう．

12.1　パラメトリック分布関数の微分

(a)　1 次元パラメータの場合

パラメトリック推定，パラメトリック仮説検定，およびこれらに関連する問題を扱う際に，パラメータ θ に依存する分布関数の導関数に関するある種の性質が必要になる．ここで簡単にこれらの問題について議論しよう．

x を c.d.f. $F(x;\theta)$ を持つ確率変数とする．ここに θ はパラメータ空間 Ω に値を持つ（実）1 次元パラメータである．ここでは便宜上，x は 1 次元確率変数とみなすが，x が k 次元の場合であっても記号を変えるだけですべての結果は成り立つ．

いま，Ω におけるいくつかの点 $\theta_1,\theta_2,\cdots$ に対応する c.d.f. $F(x;\theta_1), F(x;\theta_2),\cdots$ の集合を考える．θ の真値 θ_0 を含む Ω のある開区間 Ω_0 が特に重要である．ここで次式の両辺が θ に関して 1 回以上微分可能であると仮定する．

(12.1.1)
$$\int_{-\infty}^{\infty} dF(x;\theta)=1.$$

1 次および 2 次の微分がどうなるか考えよう．これは形式的に微分すれば，

(12.1.2)
$$\frac{d}{d\theta}\int_{-\infty}^{\infty} dF(x;\theta)=\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta}\log dF(x;\theta)\right]dF(x;\theta)=0$$

(12.1.3)
$$\frac{d^2}{d\theta^2}\int_{-\infty}^{\infty} dF(x;\theta)=\int_{-\infty}^{\infty}\left[\frac{\partial^2}{\partial\theta^2}\log dF(x;\theta)\right]dF(x;\theta)$$
$$+\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta}\log dF(x;\theta)\right]^2 dF(x;\theta)=0.$$

(12.1.2)，(12.1.3) をもう少し厳密に調べ，さらに後の参考のために，まず

(12.1.4)
$$S(x;\theta)=\frac{\partial}{\partial\theta}\log dF(x;\theta)$$

(12.1.5)
$$S'(x;\theta)=\frac{\partial}{\partial\theta}S(x;\theta)$$

(12.1.6)
$$H(\theta,\theta')=\int_{-\infty}^{\infty}\log dF(x;\theta')\,dF(x;\theta)$$

(12.1.7)
$$A(\theta,\theta')=\int_{-\infty}^{\infty}S(x;\theta')\,dF(x;\theta)$$

(12.1.8)
$$B^2(\theta,\theta')=\int_{-\infty}^{\infty}[S(x;\theta')]^2\,dF(x;\theta)$$

$$D(\theta,\theta')=\int_{-\infty}^{\infty}S'(x\,;\theta')\,dF(x\,;\theta) \tag{12.1.9}$$

と書く．ここで θ は Ω_0 の任意の点であり，(θ,θ') は積集合 $\Omega_0\times\Omega_0$ の任意の点である．

まず $S(x\,;\theta)$ と $S'(x\,;\theta)$ を考えよう．Ω_0 の任意の点 θ と確率 0 の集合をのぞくすべての x に対して，$F(x\,;\theta)$ が θ に関する 1 次導関数を持つと仮定すると，$S(x\,;\theta)$ は次式の極限が存在するとき

$$S(x'\,;\theta)=\lim_{x\to x'}\frac{\dfrac{\partial}{\partial\theta}\Big[F(x'\,;\theta)-F(x\,;\theta)\Big]}{[F(x'\,;\theta)-F(x\,;\theta)]}. \tag{12.1.10}$$

で定義される．ただし，$x<x'$ である．ここでは Ω_0 のすべての θ と R_1 のすべての x に対して極限が存在し，極限は確率正の x の値の集合では 0 でないと仮定する．$S'(x\,;\theta)$ も同様に定義される．$\theta=\theta'$ に対して，x が c.d.f. $F(x\,;\theta')$ を持つ確率変数であれば，$S(x\,;\theta)$，$S'(x\,;\theta)$ は $\Omega_0\ni\theta$ に対して確率変数になる．

x が p.f. $p(x\,;\theta)$ を持つ離散型確率変数ならば，

$$S(x\,;\theta)=\frac{\partial}{\partial\theta}\log p(x\,;\theta) \tag{12.1.10 a}$$

で，x が p.d.f. $f(x\,;\theta)$ を持つ連続型確率変数ならば

$$S(x\,;\theta)=\frac{\partial}{\partial\theta}\log f(x\,;\theta) \tag{12.1.10 b}$$

である．同じ事柄が $S'(x\,;\theta)$ に関しても離散型と連続型確率変数の場合に成り立つ．

次に，$H(\theta,\theta')$ を考えよう．x 軸を互いに素な区間 $(x_\alpha,x_{\alpha+1}]$，$\alpha=\cdots,-1,0,+1,\cdots$ に分割し，I_α で区間 $(x_\alpha,x_{\alpha+1}]$ を示す．このとき

$$P(x\in I_\alpha|\theta)=F(x_{\alpha+1}\,;\theta)-F(x_\alpha\,;\theta)$$

であり，$P(x\in I_\alpha|\theta')$ も同様な意味を持つ．$\Delta=\max\limits_{\alpha}\{I_\alpha\ の長さ\}$ とし

$$\begin{aligned} H_\Delta(\theta,\theta') &= \sum_{\alpha=-\infty}^{\infty}\log P(x\in I_\alpha|\theta')\cdot P(x\in I_\alpha|\theta) \\ &= \log\left\{\prod_{\alpha=-\infty}^{\infty}[P(x\in I_\alpha|\theta')]^{P(x\in I_\alpha|\theta)}\right\} \end{aligned} \tag{12.1.11}$$

とする．(12.1.11) の上の式のどの項も負だから，$H_\Delta(\theta,\theta')$ は負である．$-\infty$ に等しくなる項をのぞくためには，$P(x\in E|\theta)$，$P(x\in E|\theta')$ のどちらもともに 0 でもなく，

ともに正でもないような x 軸上の集合 E が存在しないと仮定すれば十分である．この条件を満足する c.d.f. $F(x\,;\theta)$，$F(x\,;\theta')$ を持つ 2 つの分布は互いに絶対連続であるという．この仮定はここで必要な条件よりも少し強い．なぜなら，$F(x\,;\theta')$ が $F(x\,;\theta)$ に関して絶対連続，すなわち，$P(x\in E|\theta')=0$ かつ $P(x\in E|\theta)>0$ なる E が存在しなければ十分である．事実，このとき $F(x\,;\theta)$ と $F(x\,;\theta')$ を入れ換えれば，互いに絶対連続になる．

$p_1,\cdots,p_r$ と $q_1,\cdots,q_r$ が正の数からなる任意の 2 組の集合のとき

$$p_1^{q_1}\cdots p_r^{q_r}\leqslant(p_1+\cdots+p_r)^{q_1+\cdots+q_r} \tag{12.1.12}$$

が成り立つので，(12.1.11) の { } の正の量は $\{I_\alpha\}$ における区間を逐次，部分分割しても増加しない．よって $H_\Delta(\theta,\theta')$ が有限になるような集合 $\{I_\alpha\}$ が存在すれば，有限な負の量 $H_\Delta(\theta,\theta')$ は $\Delta\longrightarrow0$ のとき減少できない．したがって，$\lim\limits_{\Delta\to0}H_\Delta(\theta,\theta')$ は $\Omega_0\times\Omega_0$ のすべての点 (θ,θ') で存在して正ではない．極限は (12.1.6) の積分で表わされる．$H(\theta,\theta')$ は p.f. $p(x\,;\theta)$ を持つ離散型確率変数のときは，$\sum\log p(x\,;\theta')p(x\,;\theta)$ で，p.d.f. $f(x;\theta)$ を持つ連続型確率変数のときは $\int_{-\infty}^{\infty}\log f(x\,;\theta')\,f(x\,;\theta')\,dx$ である．

関数 $H(\theta,\theta')$ は情報理論，統計的機構のエントロピー，統計的推測論におけるパラメータの最適推定，最適検定等に重要な関連を持っている．

$F(x\,;\theta)$ に関して可測な $g(x\,;\theta')$ に対して，$F(x,\theta)$ に関して可測でしかも $F(x,\theta)$ に関して有限な平均値を持つ非負関数 $h(x)$ が存在し，$(\theta,\theta')\in\Omega_0\times\Omega_0$ において

$$|g(x\,;\theta')|<h(x)$$

ならば，$g(x\,;\theta')$ は<u>積分可能な関数 $h(x)$ に支配されている</u>という．

(12.1.2) が意味を持つには，$\partial/\partial\theta'\log dF(x\,;\theta')$ がある積分可能な関数 $h_1(x)$ に支配されていれば十分である．同様に，(12.1.3) に対しても $\partial^2/\partial\theta'^2\log dF(x\,;\theta')$，$[\partial/\partial\theta'\log dF(x\,;\theta')]^2$ がある積分可能な関数 $h_2(x)$，$h_3(x)$ で支配されていればよい．このような十分条件に関する一般定理については，McShane (1944)，Mc-Shane と Botts (1959)，Saks (1937) による積分論と実変数解析論を参照せよ．

(12.1.2), (12.1.3) を平均値で表わして

$$\mathscr{E}(S(x\,;\theta))=\frac{d}{d\theta}\int_{-\infty}^{\infty}dF(x\,;\theta)=0 \tag{12.1.2 a}$$

ならば，$F(x\,;\theta)$ は Ω_0 において θ の **1** 次導関数に関して正則で，また $B^2(\theta,\theta)<+\infty$ でかつ

(12.1.3 a)　$\mathscr{E}(S'(x;\theta))+\mathscr{E}(S(x;\theta))^2=\dfrac{d^2}{d\theta^2}\displaystyle\int_{-\infty}^{\infty}dF(x;\theta)=0$

ならば，Ω_0 において θ の **2 次導関数に関して正則である**という．

ここで (12.1.6) で定義された $H(\theta,\theta')$ に戻ろう．$F(x;\theta)$ が θ の最初の 2 つの導関数に関して正則ならば，$H(\theta,\theta')$ は $\Omega_0\times\Omega_0$ 上で θ と θ' に関する 1 次偏導関数を持つ．さらに θ を固定すると θ' の関数として $H(\theta,\theta')$ は，$\theta'=\theta$ のとき最大値を持つ．

(b)　ベクトルパラメータの場合

さて，$F(x;\theta)$ を c.d.f. としよう．ただし θ のパラメータ空間は Ω_r で，θ は（関数的に独立な）成分 $(\theta_1,\cdots,\theta_r)$ を持つ r 次元パラメータである．便宜上，x を 1 次元とする．しかし，以下の結果は表現を変えれば，k 次元の場合にもそのまま成り立つ．

$$\int_{-\infty}^{\infty}dF(x;\theta)$$

が積分記号下で θ のどの成分についても何回でも微分可能になるような十分条件の下では，以下の論議には特に困難はないだろう．2 つの典型的な微分だけを考えると，(12.1.2) と (12.1.3) に対応して

(12.1.13)　$\dfrac{\partial}{\partial\theta_p}\displaystyle\int_{-\infty}^{\infty}dF(x;\theta)=\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta_p}\log dF(x;\theta)\right]dF(x;\theta)=0$

(12.1.14)　$\dfrac{\partial^2}{\partial\theta_p\partial\theta_q}\displaystyle\int_{-\infty}^{\infty}dF(x;\theta)=\int_{-\infty}^{\infty}\left[\frac{\partial^2}{\partial\theta_p\partial\theta_q}\log dF(x;\theta)\right]dF(x;\theta)$

$$+\int_{-\infty}^{\infty}\left[\frac{\partial}{\partial\theta_p}\log dF(x;\theta)\right]\left[\frac{\partial}{\partial\theta_q}\log dF(x;\theta)\right]dF(x;\theta)=0$$

が得られる．

$p,q=1,\cdots,r$ および r 次元（ユークリッド）開区間 Ω_{r0} のすべての点に対して (12.1.13) と (12.1.14) が成り立つかどうかに注意すべきである．

(12.1.13), (12.1.14) の証明および後の参考のために，(12.1.4) から (12.1.9) に対応させて次のように置こう．

(12.1.15)　$S_p(x;\theta)=\dfrac{\partial}{\partial\theta_p}\log dF(x;\theta)$

(12.1.16)　$S_{pq}(x;\theta)=\dfrac{\partial^2}{\partial\theta_p\partial\theta_q}\log dF(x;\theta)$

(12.1.17)　$H(\theta,\theta')=\displaystyle\int_{-\infty}^{\infty}\log dF(x;\theta')\,dF(x;\theta)$

(12.1.18)　$A_p(\theta,\theta')=\displaystyle\int_{-\infty}^{\infty}S_p(x;\theta')\,dF(x;\theta)$

(12.1.19)　$B_{pq}(\theta,\theta')=\displaystyle\int_{-\infty}^{\infty}S_p(x;\theta')S_q(x;\theta')\,dF(x;\theta)$

(12.1.20)　$D_{pq}(\theta,\theta')=\displaystyle\int_{-\infty}^{\infty}S_{pq}(x;\theta')\,dF(x;\theta)$

ここで $p,q=1,\cdots,r$，θ は Ω_{r0} の点で，(θ,θ') は $\Omega_{r0}\times\Omega_{r0}$ の点である．θ の成分は関数的に独立であると仮定してあるので，確率変数 $(S_p(x;\theta),p=1,\cdots,r)$ の成分は 1 次独立である．よって行列 $\|B_{pq}(\theta,\theta')\|$ は $\Omega_{r0}\times\Omega_{r0}$ の (θ,θ') に対して正定値である．1 次元パラメータの議論と同様であるから，これらの関数には特に注意はつけ加えないでもよかろう．

1 次元パラメータの場合から，(12.1.13) が成り立つのは，$(\partial/\partial\theta'_p)\log dF(x;\theta')$ が積分可能な関数 $h_{1p}(x)$，$p=1,\cdots,r$ で支配されていれば十分である．同様に，(12.1.14) が成り立つためには

$$(\partial^2/\partial\theta'_p\partial\theta'_q)\log dF(x;\theta')$$

および

$$[(\partial/\partial\theta'_p)\log dF(x;\theta')][(\partial/\partial\theta'_q)\log dF(x;\theta')]$$

がそれぞれ積分可能な関数 $h_{2pq}(x)$，$h_{3pq}(x)$，$p,q=1,\cdots,r$ で支配されていればよい．

(12.1.13), (12.1.14) を平均値で表わして，

(12.1.13 a)　$\mathscr{E}(S_p(x;\theta))=\dfrac{\partial}{\partial\theta_p}\displaystyle\int_{-\infty}^{\infty}dF(x;\theta)=0,\qquad p=1,2,\cdots,r$

ならば，$F(x;\theta)$ は Ω_{r0} においてその θ の **1 次偏導関数に関して正則である**という．

また $\|B_{pq}(\theta,\theta')\|$ が有限で

(12.1.14 a)　$\mathscr{E}(S_p(x;\theta)S_q(x;\theta))+\mathscr{E}(S_{pq}(x;\theta))=\dfrac{\partial^2}{\partial\theta_p\partial\theta_q}\displaystyle\int_{-\infty}^{\infty}dF(x;\theta)=0$

ならば，$F(x;\theta)$ は Ω_{r0} においてその θ の **2 次偏導関数に関して正則である**という．

(c)　ベクトル確率変数への拡張に関する注意

これまでの議論は 1 次元確率変数の場合であった．若干の変更で k 次元確率変数の場合に議論が展開できる．x がベクトル $(x_1,\cdots,x_k)$ のとき，問題は $S(x;\theta)$，$H(\theta,\theta')$ の定義である．この場合，(12.1.10) での区間 $(x,x']$ 上の $F(x;\theta)$ の差分 $[F(x';\theta)$

$-F(x;\theta)]$ は k 次元区間 $(x_1,\cdots,x_k;x_1',\cdots,x_k']$ 上の $F(x_1,\cdots,x_k;\theta)$ の差分に置き換えられる．極限の存在を仮定すれば，$(\partial/\partial\theta)\log dF(x_1,\cdots,x_k;\theta)$ が定義され，$S(x_1,\cdots,x_k;\theta)$ で表わされる．

同様に，(12.1.11) における $(x_\alpha,x_{\alpha+1}]$ 上の $F(x;\theta)$ の差分 $F(x_{\alpha+1};\theta)-F(x_\alpha;\theta)$ は k 次元区間 $(x_{1,\alpha},\cdots,x_{k,\alpha};x_{1,\alpha+1},\cdots,x_{k,\alpha+1}]$ 上での $F(x_1,\cdots,x_k;\theta)$ の k 次元差分に置き換えられる．ただし $\Delta=\max_\alpha\{\Delta_\alpha\}$．$\Delta_\alpha$ はこの k 次元区間の最大の長さである．

記号の変換は明らかであろう．

12.2　点　推　定

(a)　定　　義

$(x_1,\cdots,x_n)$ を c.d.f. $F_n(x_1,\cdots,x_n;\theta)$ からの n 次元確率変数とする．ここで θ はパラメータ空間 Ω の 1 次元実数パラメータとする．

$\tilde\theta(x_1,\cdots,x_n)$ あるいは簡単に $\tilde\theta$ を $\tilde\theta$ 自身が確率変数である $(x_1,\cdots,x_n)$ の関数とする．$(x_1,\cdots,x_n)$ の実現（観測）値に対応する $\tilde\theta$ の実現（観測）値が θ の真値 θ_0 の推定に使用されるとき，確率変数 $\tilde\theta$ は θ_0 の点推定または推定量といわれる．もちろん θ_0 が未知のときにのみ $\tilde\theta$ は使用される．$\theta=\theta_0$ のとき $\mathscr{E}(\tilde\theta)=\theta_0$（簡単に $\mathscr{E}(\tilde\theta|\theta_0)=\theta_0$ と書く）ならば，$\tilde\theta$ を θ_0 に対する不偏推定量という．正確にいえば，$\tilde\theta$ は θ_0 に対して平均において偏りのない推定量である．$\tilde\theta$ が，$\tilde\theta=\theta_0$ で連続でしかも $W(\theta_0;\theta_0)=\frac{1}{2}$ となるような c.d.f. $W(\tilde\theta;\theta_0)$ を持つ統計量ならば，$\tilde\theta$ はメディアンにおいて偏りのない θ_0 に対する推定量という．しかし特に，明示しないかぎり不偏推定量は平均において不偏であると解釈してよい．

推定量 $\tilde\theta$ が，$n\longrightarrow\infty$ のとき，θ_0 に確率収束するならば $\tilde\theta$ は θ_0 に対する**一致推定量**という．

$\tilde\theta$ が分散有限な θ_0 の不偏推定量で，他により小さい分散を持つ推定量が存在しないとき，$\tilde\theta$ は θ_0 に対する**有効推定量**という．

$\tilde\theta$ が他の任意の統計量 $\check\theta$ に対して，条件つき確率変数 $\check\theta|\tilde\theta$ の分布が θ_0 に依存しないような統計量であるとき，$\tilde\theta$ は θ_0 に対する**十分統計量**という．さらに $\mathscr{E}(\tilde\theta|\theta_0)=\theta_0$ のとき $\tilde\theta$ は θ_0 の**十分推定量**という．

一致推定量，有効推定量，十分推定量の概念は Fisher (1922, 1925 *b*) による．

単純無作為標本のとき，すなわち $(x_1,\cdots,x_n)$ が c.d.f. $F(x;\theta)$ からの大きさ n の確率標本のとき，不偏性，一致性，有効性，十分性の概念は特に重要である．この場合

(12.2.1)　$$F_n(x_1,\cdots,x_n;\theta)=\prod_{\xi=1}^{n}F(x_\xi;\theta)$$

となる．与えられた標本 $(x_1,\cdots,x_n)$ に対して，$dF_n=\prod_{\xi=1}^{n}dF(x_\xi;\theta)$ を $(x_1,\cdots,x_n)$ に対する θ の**尤度要素***) という．

本章では主に単純無作為標本に関するものを取りあげる．

(b)　推定量の分散の下界

$(x_1,\cdots,x_n)$ を c.d.f. $F_n(x_1,\cdots,x_n;\theta)$ を持つ n 次元確率変数とする．$\tilde\theta(x_1,\cdots,x_n)$ が θ の不偏推定量ならば，

(12.2.2)　$$\int_{R_n}(\tilde\theta-\theta)\,dF_n=0.$$

$F_n(x_1,\cdots,x_n;\theta)$ が θ の真値 θ_0 を含むある（開）区間 Ω_0 においてその θ の 1 次導関数に関して正則ならば，積分記号下で (12.2.2) を微分すると

(12.2.3)　$$\int_{R_n}(\tilde\theta-\theta)S_n(x_1,\cdots,x_n;\theta)\,dF_n=1$$

を得る．ただし，

(12.2.4)　$$S_n(x_1,\cdots,x_n;\theta)=\frac{\partial\log dF_n}{\partial\theta}$$

である．

シュワルツの不等式を (12.2.3) に用いると，

$$1=\left[\int_{R_n}(\tilde\theta-\theta)S_n\,dF_n\right]^2\leqslant\int_{R_n}(\tilde\theta-\theta)^2\,dF_n\cdot\int_{R_n}S_n^2\,dF_n.$$

しかし，

(12.2.5)　$$\sigma^2(\tilde\theta)=\int_{R_n}(\tilde\theta-\theta)^2\,dF_n$$

より

(12.2.6)　$$\sigma^2(\tilde\theta)\geqslant\frac{1}{\mathscr{E}(S_n^2)},\qquad\theta_0\in\Omega_0$$

が得られる．等号が成り立つための必要十分条件は確率 1 で R_n において

*) 確率素分に対応して，尤度素分といってもよい．[訳注]

(12.2.7)
$$K[\tilde{\theta}(x_1,\cdots,x_n)-\theta]\equiv S_n(x_1,\cdots,x_n\,;\,\theta)$$

が成り立つことである．ただし K は $(x_1,\cdots,x_n)$ ではなく，θ に依存する．

(12.2.6) 式で等号が成り立つとき，$\tilde{\theta}$ を θ に対する **有効推定量** といい，この場合，(12.2.7) で $\tilde{\theta}$ の型が与えられる．

$\theta=\theta_0$ のとき，すなわち，(12.2.5) で $\theta=\theta_0$ と置いたとき，$\sigma^2(\tilde{\theta})$ を $\sigma^2(\tilde{\theta}|\theta_0)$ で表わし，$\mathscr{E}(S_n^2|\theta_0)$ も同じ意味で用いれば

$$\sigma^2(\tilde{\theta}|\theta_0)\geqslant 1/[\mathscr{E}(S_n^2|\theta_0)]$$

となる．θ の有効推定量は通常 $\hat{\theta}$ で表わされる．まとめると，

12.2.1　$(x_1,\cdots,x_n)$ を c.d.f. $F_n(x_1,\cdots,x_n\,;\,\theta)$ を持つ確率変数とする．ただし θ は 1 次元で，$F_n(x_1,\cdots,x_n\,;\,\theta)$ は Ω_0 においてその θ の 1 次導関数に関して正則であるとする．$\tilde{\theta}(x_1,\cdots,x_n)$ が θ_0 の不偏推定量ならば，$\sigma^2(\tilde{\theta}|\theta_0)\geqslant 1/[\mathscr{E}(S_n^2|\theta_0)]$ である．ここで S_n は (12.2.4) で与えられ，等号が成り立つための必要十分条件は $\theta=\theta_0$ で (12.2.7) が確率 1 で成り立つことである．θ_0 に対する有効推定量 $\hat{\theta}$ が存在すればその分散は $1/\mathscr{E}(S_n^2|\theta_0)$ である．

正則性を仮定しないとき，$\sigma^2(\tilde{\theta}|\theta_0)$ に対する下界は Chapman と Robbins (1951), Kiefer (1952) が与えている．

θ_0 に対する有効推定量 $\hat{\theta}$ が存在して，$\tilde{\theta}$ が θ_0 の他の任意の不偏推定量ならば，$\tilde{\theta}$ の θ_0 の推定量に対する効率は比

(12.2.8)
$$\mathrm{eff}\,(\tilde{\theta}|\theta_0)=\frac{\sigma^2(\hat{\theta}|\theta_0)}{\sigma^2(\tilde{\theta}|\theta_0)}$$

で与えられる．

(c)　無作為標本の場合

この場合 (12.2.1) は成り立ち，

(12.2.9)
$$S_n(x_1,\cdots,x_n\,;\,\theta)=\sum_{\xi=1}^{n}S(x_\xi\,;\,\theta)$$

である．ただし

(12.2.10)
$$S(x\,;\,\theta)=\frac{\partial}{\partial\theta}\log dF(x\,;\,\theta)$$

$\sum_{\xi=1}^{n}S(x_\xi\,;\,\theta)$ を $(x_1,\cdots,x_n)$ に基づく θ に対する**スコア**という．

さらに

(12.2.11)
$$\mathscr{E}(S_n^2)=n\mathscr{E}(S^2)=nB^2(\theta,\theta)$$

が成り立つ．ただし $B^2(\theta,\theta)$ は (12.1.8) で与えられている．それゆえ (12.2.6) は

(12.2.12)
$$\sigma^2(\tilde{\theta})\geqslant\frac{1}{nB^2(\theta,\theta)}$$

となる．この結果は Fisher (1922) により最初に示された．その後 Cramér (1946), Dugué (1937), Rao (1945) にも発表されている．

(12.2.12) の等号が成り立つための必要十分条件は R_n 上で

(12.2.13)
$$K(\tilde{\theta}-\theta)\equiv\sum_{\xi=1}^{n}S(x_\xi\,;\,\theta)$$

が確率 1 で成り立つことである．ただし K は $(x_1,\cdots,x_n)$ に依存しない．よって有効推定量 $\hat{\theta}$ が存在すれば，それは (12.2.13) を満足する統計量 $\tilde{\theta}$ であり，その分散は

(12.2.14)
$$\sigma^2(\hat{\theta})=\frac{1}{nB^2(\theta,\theta)}$$

で与えられる．この結果，次の **12.2.1** の重要な系を得る．

12.2.1 a　$(x_1,\cdots,x_n)$ が Ω_0 でその θ の 1 次導関数に関して正則な c.d.f. $F(x\,;\,\theta)$ からの標本で，$\tilde{\theta}$ が θ_0 の不偏推定量であれば，$\sigma^2(\tilde{\theta}|\theta_0)\geqslant 1/[nB^2(\theta_0,\theta_0)]$ である．ただし $B^2(\theta,\theta)$ は (12.1.8) で与えられる．さらに等号が成り立つための必要十分条件は $\tilde{\theta}$ が (12.2.13) を満足することである．この場合，$\tilde{\theta}$ に対する解（これを $\hat{\theta}$ で表わす）は分散 $1/[nB^2(\theta_0,\theta_0)]$ の θ_0 に対する推定量である．

$\tilde{\theta}$ が θ に対する任意の不偏推定量で，θ の有効推定量 $\hat{\theta}$ が存在すれば，θ_0 を推定する $\tilde{\theta}$ の効率は

(12.2.15)
$$\mathrm{eff}\,(\tilde{\theta}|\theta_0)=\frac{\sigma^2(\hat{\theta}|\theta_0)}{\sigma^2(\tilde{\theta}|\theta_0)}=\frac{1}{\sigma^2(\sqrt{n}\,\tilde{\theta}|\theta_0)B^2(\theta_0,\theta_0)}$$

で定義される．

Fisher (1922) は $nB^2(\theta_0,\theta_0)$ ——有効推定量 $\hat{\theta}$ の分散の逆数——を θ_0 に関する標本に含まれる**情報量**，あるいは $B^2(\theta_0,\theta_0)$ を $F(x\,;\,\theta_0)$ からの観測あたりの θ_0 に関する情

報量と呼んでいる.

この観点から, θ に対する不偏推定量 $\tilde{\theta}$ の効率は, θ_0 を推定する有効推定量に含まれる情報に対する, θ_0 を推定する $\tilde{\theta}$ に含まれる情報の割合であるといえる.

例題 $(x_1, \cdots, x_n)$ は **8.3.3c** で示したポアソン分布 $Po(\mu_0)$ からの標本とする. このとき,

$$S_n(x_1, \cdots, x_n\,;\,\mu) = \frac{\partial}{\partial\mu}\log\prod_{\xi=1}^{n} dF(x_\xi\,;\,\mu) = \frac{n}{\mu}\left[\frac{1}{n}\sum_{\xi=1}^{n} x_\xi - \mu\right]$$

を得る. (12.2.13) に照らして, これは標本の大きさが任意に与えられたとき, 標本平均 $\frac{1}{n}\sum_{\xi=1}^{n} x_\xi$ が μ の真値 μ_0 の有効推定量であることを示している. $\frac{1}{n}\sum_{\xi=1}^{n} x_\xi$ を $\hat{\mu}(x_1, \cdots, x_n)$ あるいは $\hat{\mu}$ で表わせば, (12.2.14) を用いると $\hat{\mu}$ の分散が求まる. すなわち

$$\sigma^2(\hat{\mu}|\mu_0) = \frac{1}{n\sum_{x=0}^{\infty}\left(\frac{x}{\mu_0} - 1\right)^2 \frac{\mu_0^x e^{-\mu_0}}{x!}} = \frac{\mu_0}{n}.$$

また

$$\sigma^2(\sqrt{n}\hat{\mu}|\mu_0) = \mu_0$$

となり, **12.2.1a** から $\sigma^2(\sqrt{n}\tilde{\mu}|\mu_0) < \mu_0$ となる μ_0 の不偏推定量 $\tilde{\mu}(x_1, \cdots, x_n)$ を求めるのは不可能なことがわかる.

(d) 偏りのある推定量の分散の下界

12.2.1 の一般化 [Cramér (1946), Dugué (1937), Rao (1945) を参照] は次のように θ に対する偏りのある推定量 $\tilde{\theta}$ を考えることによって得られる. (12.2.2) の $(\tilde{\theta} - \theta)$ を

(12.2.16) $$[\tilde{\theta} - \theta - b_n(\theta)]$$

で置き換える. ただし, $b_n(\theta)$ は推定量 $\tilde{\theta}$ の偏りである. こうして **12.2.1a** の場合と本質的に同じ議論を繰り返す. このとき, (12.2.3) の $(\tilde{\theta} - \theta)$ は $(\tilde{\theta} - \theta - b_n(\theta))$ で置き換え, 右辺の 1 は $1 + b'_n(\theta)$ で置き換える. ただし $b'_n(\theta) = d/d\theta b_n(\theta)$. (12.2.12) にかわって

(12.2.17) $$\sigma^2(\tilde{\theta}) \geqq \frac{[1 + b'_n(\theta)]^2}{nB^2(\theta, \theta)}$$

が得られる. 等号が成り立つための必要十分条件は R_n 上で

(12.2.18) $$K[\tilde{\theta} - \theta - b_n(\theta)] \equiv \sum_{\xi=1}^{n} S(x_\xi\,;\,\theta)$$

が確率 1 で成り立つときである. ただし K は $(x_1, \cdots, x_n)$ に依存しない.

(e) 十分推定量の性質

$(x_1, \cdots, x_n)$ を p.d.f. $f_n(x_1, \cdots, x_n\,;\,\theta)$ を持つ確率変数とし, $f_n(x_1, \cdots, x_n\,;\,\theta)$ は次のように因数分解されるとする.

(12.2.19) $$f_n(x_1, \cdots, x_n\,;\,\theta) = v(\tilde{\theta}\,;\,\theta)u(x_1, \cdots, x_n|\tilde{\theta}).$$

ここで $v(\tilde{\theta}\,;\,\theta)$ は $\tilde{\theta}$ の p.d.f. で, $u(x_1, \cdots, x_n|\tilde{\theta})$ は θ に依存しない条件つき確率変数 $(x_1, \cdots, x_n|\tilde{\theta})$ の p.d.f. である. このとき $\tilde{\theta}$ は θ の十分統計量である. なぜなら $\breve{\theta}$ が $\tilde{\theta}$ に依存しない他の任意の統計量ならば, 条件つき確率変数 $\breve{\theta}|\tilde{\theta}$ の分布は $u(x_1, \cdots, x_n|\tilde{\theta})$ により完全に定まるからである.

逆に, $\tilde{\theta}$ を p.d.f. $v(\tilde{\theta}\,;\,\theta)$ を持つ十分統計量とし, さらに $y_2, \cdots, y_n$ を $\tilde{\theta}$, $y_2, \cdots, y_n$ が次の p.d.f. を持つ $n-1$ 個の任意の統計量とする.

(12.2.20) $$g(\tilde{\theta}, y_2, \cdots, y_n\,;\,\theta) = f_n(x_1, \cdots, x_n\,;\,\theta)\frac{1}{|J|}.$$

ここで J は $(x_1, \cdots, x_n)$ に関する $(\tilde{\theta}, y_2, \cdots, y_n)$ のヤコビアンである [(12.2.20) が成り立つための条件は 2.8(d) 節を見よ].

$h(y_2, \cdots, y_n|\tilde{\theta}\,;\,\theta)$ を条件つき確率変数 $y_2, \cdots, y_n|\tilde{\theta}$ の p.d.f. とすれば

(12.2.21) $$h(y_2, \cdots, y_n|\tilde{\theta}\,;\,\theta) = \frac{g(\tilde{\theta}, y_2, \cdots, y_n\,;\,\theta)}{v(\tilde{\theta}\,;\,\theta)}$$

である. ただし $v(\tilde{\theta}\,;\,\theta) \neq 0$ であり,

(12.2.22) $$v(\tilde{\theta}\,;\,\theta) = \int_{R_{n-1}} g(\tilde{\theta}, y_2, \cdots, y_n\,;\,\theta)\,dy_2 \cdots dy_n$$

により与えられる.

$h(y_2, \cdots, y_n|\tilde{\theta}\,;\,\theta)$ が θ に依存しない (このとき, $h^*(y_2, \cdots, y_n|\tilde{\theta})$ で表わす) ならば, $\tilde{\theta}$ は θ の十分統計量で

(12.2.23) $$g(\tilde{\theta}, y_2, \cdots, y_n\,;\,\theta) = v(\tilde{\theta}\,;\,\theta)\,h^*(y_2, \cdots, y_n|\tilde{\theta})$$

となる. (12.2.20) を用いると

(12.2.24) $$f_n(x_1, \cdots, x_n\,;\,\theta) = v(\tilde{\theta}\,;\,\theta)h^*(y_2, \cdots, y_n|\tilde{\theta})\cdot|J|$$

が得られる.

$h^*(y_2, \cdots, y_n|\tilde{\theta})\cdot|J|$ は $(\tilde{\theta}, y_2, \cdots, y_n)$, したがって $(x_1, \cdots, x_n)$ に依存する. しかし, θ には依存しない. よって $(y_2, \cdots, y_n)$ を通じて $(x_1, \cdots, x_n)$ に依存するが θ には依存

しない任意の統計量は θ に依存しない分布を持つ.

要約すると

12.2.2　$(x_1,\cdots,x_n)$ を p.d.f. $f_n(x_1,\cdots,x_n;\theta)$ を持つ確率変数とする. 統計量 $\tilde{\theta}(x_1,\cdots,x_n)$ が θ の十分統計量であるための必要十分条件は,

(12.2.25) $$f_n(x_1,\cdots,x_n;\theta)=v(\tilde{\theta};\theta)w(x_1,\cdots,x_n)$$

である. ここで $v(\tilde{\theta};\theta)$ は $\tilde{\theta}$ の p.d.f.で, $w(x_1,\cdots,x_n)$ は θ に依存しない $(x_1,\cdots,x_n)$ の関数である.

Fisher (1922) は分解基準式 (12.2.25) の十分性を最初に指摘し, Neyman (1935) はこれが必要条件であることを示した.

$(x_1,\cdots,x_n)$ が p.d.f. $f(x;\theta)$ からの標本であるような単純無作為標本の特殊ではあるが重要な場合には, $f_n(x_1,\cdots,x_n;\theta)$ が $\prod_{\xi=1}^{n}f(x_\xi;\theta)$ なる特殊形をとる. このとき **12.2.2** の系が得られる.

同様な方法で, $(x_1,\cdots,x_n)$ が p.d.f. $p_n(x_1,\cdots,x_n;\theta)$ を持つ n 次元離散型確率変数のとき, $\tilde{\theta}$ が θ の十分統計量であるための必要十分条件は $p_n(x_1,\cdots,x_n;\theta)$ が

(12.2.26) $$p_n(x_1,\cdots,x_n;\theta)=v(\tilde{\theta};\theta)\,w(x_1,\cdots,x_n)$$

と因数分解できることであることがわかる. ただし, $v(\tilde{\theta};\theta)$ は $\tilde{\theta}$ の p.f.で, $w(x_1,\cdots,x_n)$ は $(x_1,\cdots,x_n)$ に依存するが θ には依存しない.

p.f. $p(x;\theta)$ からの単純無作為標本では, $p_n(x_1,\cdots,x_n;\theta)$ は $\prod_{\xi=1}^{n}p(x_\xi;\theta)$ なる型をとる.

最後に, **12.2.2** は θ がベクトル $(\theta_1,\cdots,\theta_r)$, $r\leqslant n$ の場合にも拡張できることを示そう. $(\tilde{\theta}_1,\cdots,\tilde{\theta}_s)$ が $(\theta_1,\cdots,\theta_r)$, $s\geqslant r$ の十分統計量の組であるための必要十分条件は

(12.2.27) $$f_n(x_1,\cdots,x_n;\theta_1,\cdots,\theta_r)=v(\tilde{\theta}_1,\cdots,\tilde{\theta}_s;\theta_1,\cdots,\theta_r)\,w(x_1,\cdots,x_n)$$

である. ただし $w(x_1,\cdots,x_n)$ は $(x_1,\cdots,x_n)$ には依存するが θ には依存しない.

θ がベクトル $(\theta_1,\cdots,\theta_r)$ で, $(x_1,\cdots,x_n)$ が p.f. $p_n(x_1,\cdots,x_n;\theta_1,\cdots,\theta_r)$ を持つ離散型確率変数の場合にも分解基準が適用される.

Halmos と Savage (1949) は一般の分布を持つ確率変数 $(x_1,\cdots,x_n)$ に対して, ベクトル値パラメータの場合へ拡張するとともに, **12.2.2** の分解基準をラドン＝ニコディムの定理を使用して一般化した. また, 十分統計量を特徴づけるための必要十分条件の抽象的な取り扱いは Bahadur (1954) によりなされた.

例題　十分推定量の例として, $(x_1,\cdots,x_n)$ をポアソン分布 $Po(\mu_0)$ からの標本と仮定する. このとき, $(0,\infty)$ 上の任意の μ に対して

$$dF(x;\mu)=\frac{\mu^x e^{-\mu}}{x\,!}$$

かつ

$$\prod_{\xi=1}^{n}dF(x_\xi;\mu)=\frac{\mu^{\sum_\xi x_\xi}e^{-n\mu}}{\prod_\xi x_\xi\,!}$$

となる. これは

$$\left[\frac{\mu^{\sum_\xi x_\xi}e^{-n\mu}}{\left(\sum_\xi x_\xi\right)!}\right]\cdot\left\{\frac{\left(\sum_\xi x_\xi\right)!}{\prod_\xi x_\xi\,!}\right\}$$

と書ける. $\frac{1}{n}\sum_\xi x_\xi$ を $\tilde{\mu}(x_1,\cdots,x_n)$ として, **8.3.3c** を参照すると [ ] 内は $\tilde{\mu}$ の p.f. である. すなわち

$$dV(\tilde{\mu},\mu)=\frac{\mu^{n\tilde{\mu}}e^{-n\mu}}{(n\tilde{\mu})\,!}.$$

一方, { } 内は, 事実, 条件つき確率変数 $(x_1,\cdots,x_n|\tilde{\mu})$ の p.f. となる. すなわち

$$dU(x_1,\cdots,x_n|\tilde{\mu})=\frac{(n\tilde{\mu})\,!}{\prod_\xi x_\xi\,!}.$$

$\tilde{\mu}^*$ が他の任意の推定量ならば, 条件つき確率変数 $\tilde{\mu}^*|\tilde{\mu}$ の p.f. $p(\tilde{\mu}^*|\tilde{\mu})$ は 2 つの条件式 $\tilde{\mu}^*(x_1,\cdots,x_n)=\tilde{\mu}^*$, $\tilde{\mu}(x_1,\cdots,x_n)=\tilde{\mu}$ を満たす x_ξ の正（または零）整数にわたる $(n\tilde{\mu})\,!\big/\prod_\xi x_\xi\,!$ の和をとることにより容易に得られる. $p(\tilde{\mu}^*|\tilde{\mu})$ は μ に依存しないことがわかる. よって $\frac{1}{n}\sum_\xi x_\xi$ は μ_0 の十分推定量である.

十分推定量には次のような重要な性質がある. 十分推定量の関数でない, パラメータの任意の初期不偏推定量から出発したとしても, 初期推定量の分散よりも少さい分散を持つ十分推定量に依存する不偏推定量を見つけることができる. 正確には次の定理に述べられる. これは Blackwell (1947) と Rao (1945) による.

12.2.3　$\tilde{\theta}$ を θ_0 の十分統計量, $\check{\theta}$ を θ_0 の任意の不偏推定量とする. また $\mathscr{E}(\check{\theta}|\tilde{\theta})=h(\tilde{\theta})$ とする. このとき $h(\tilde{\theta})$ は θ_0 の不偏推定量であり, この分散は $\check{\theta}$ の分

散を越えない．

12.2.3 の証明には，$G(\bar{\theta},\breve{\theta};\theta_0)$ を $(\bar{\theta},\breve{\theta})$ の c.d.f.，$V(\bar{\theta};\theta_0)$ を $\bar{\theta}$ の c.d.f.，$U(\breve{\theta}|\bar{\theta})$ を条件つき確率変数 $\breve{\theta}|\bar{\theta}$ の c.d.f. とする．このとき

(12.2.28) $$h(\bar{\theta}) = \int_{-\infty}^{\infty}\breve{\theta}\,dU(\breve{\theta}|\bar{\theta})$$

となる．よって，$h(\bar{\theta})$ は θ_0 の不偏推定量である．

$\breve{\theta}$ の分散に対して

$$\begin{aligned}\sigma^2(\breve{\theta}) &= \mathscr{E}(\breve{\theta}-\theta_0)^2 = \mathscr{E}[(h(\bar{\theta})-\theta_0)+(\breve{\theta}-h(\bar{\theta}))]^2\\ &= \sigma^2(h(\bar{\theta}))+\mathscr{E}(\breve{\theta}-h(\bar{\theta}))^2+2\mathscr{E}[(h(\bar{\theta})-\theta_0)(\breve{\theta}-h(\bar{\theta}))].\end{aligned}$$

しかし

$$\begin{aligned}&\mathscr{E}[(h(\bar{\theta})-\theta_0)(\breve{\theta}-h(\bar{\theta}))]\\ &\quad=\int_{-\infty}^{\infty}[h(\bar{\theta})-\theta_0]\left\{\int_{-\infty}^{\infty}(\breve{\theta}-h(\bar{\theta}))\,dU(\breve{\theta}|\bar{\theta})\right\}dV(\bar{\theta};\theta_0)=0.\end{aligned}$$

なぜならば，{　} の中は消えてなくなる．ゆえに $\mathscr{E}(\breve{\theta}-h(\bar{\theta}))^2 \geqq 0$ であるから

(12.2.29) $$\sigma^2(\breve{\theta}) \geqq \sigma^2(h(\bar{\theta})).$$

これで **12.2.3** の議論を終わる．

12.3　大標本からの点推定

(a)　スコアの漸近分布

パラメータの有効推定量は大きさ n の標本に対して c.d.f. $F(x;\theta)$ がある特定の場合のみに存在する．このような場合，(12.2.9) で与えられた $S_n(x_1,\cdots,x_n;\theta)$ は (12.2.13) のような特殊型になり，したがって，有効推定量 $\hat{\theta}_n(x_1,\cdots,x_n)$ は本質的には次式を θ について解くことにより得られる．

(12.3.1) $$S_n(x_1,\cdots,x_n;\theta)=0.$$

しかし，$S_n(x_1,\cdots,x_n;\theta)$ が (12.2.13) で示された特殊型を持たず，(12.3.1) が θ に対する解を持つと仮定して，この解を $\hat{\theta}_n(x_1,\cdots,x_n)$ と引き続き呼ぶことにしよう．この解は θ_0 の推定量としてどのような性質を持つであろうか．これに対する答は次のようになる．以下に述べる条件の下では，$\hat{\theta}_n(x_1,\cdots,x_n)$ は大標本に対する漸近的な意味で θ_0 の有効推定量になる．

この問題を扱うに先だって次の結果を示しておこう．

12.3.1　$(x_1,\cdots,x_n)$ を c.d.f. $F(x;\theta_0)$ からの標本とする．ただし，$F(x;\theta)$ は Ω_0 で θ の 1 次導関数に関して正則とする．このとき，(12.1.8) で定義された $B^2(\theta;\theta)$ が存在して有限ならば，$S_n(x_1,\cdots,x_n;\theta_0)$ は漸近的に $N(0,nB^2(\theta_0,\theta_0))$ に従って分布する．

これは本質的には **9.2.1** の系である．なぜなら確率変数 $S(x;\theta)$ を y とすれば，y は

$$G(y)=\int_{E_y}dF(x;\theta)$$

なる c.d.f. $G(y)$ を持つからである．ここで，E_y は $S(x;\theta)\leqq y$ なる x 軸上の点集合である．$F(x;\theta)$ は θ の 1 次導関数に関して正則だから $\mathscr{E}(y)=0$．よって $S_n(x_1,\cdots,x_n;\theta)$ は平均 0，有限な分散 $B^2(\theta,\theta)$ を持つ母集団からの大きさ n の標本和である．**9.2.1** を適用すれば **12.3.1** の結論が導かれる．

(b)　最尤推定量の収束

12.3.1 の条件下では，確率変数 $(1/n)S_n(x_1,\cdots,x_n;\theta)$ は明らかに 0 に確率収束する．これは

(12.3.2) $$S_n(x_1,\cdots,x_n;\theta)=0$$

ならば，根の集合の列 $\{\hat{\theta}(x_1,\cdots,x_n)\}$，$n=1,2,\cdots$ が得られて，各集合には，θ の真値に確率 1 で収束する列が求められるような根を少なくとも 1 つ持つようにすることができることを示している．ある条件下ではこれが正しいことを示そう．$S(x;\theta)$ は確率 0 の集合をのぞいた R_1 のすべての x の値に対して Ω_0 において θ の連続関数であると仮定する．さらに，$F(x;\theta)$ は Ω_0 において θ の 1 次導関数に関して正則，したがって (12.1.7) で定義された $A(\theta_0,\theta)$ は連続かつ θ_0 を含む Ω_0 のある部分区間 Ω_0' 上で θ に関して狭義減少と仮定する．

(12.2.9) と，**12.3.1** の証明を参照すると，θ_0 が θ の真値のとき，$(1/n)S_n(x_1,\cdots,x_n;\theta)$ は明らかに平均 $A(\theta_0,\theta)$ を持つ母集団からの大きさ n の標本平均である．それゆえ，**4.6.1** より，$(1/n)S_n(x_1,\cdots,x_n;\theta)$ は $A(\theta_0,\theta)$ に概収束する．一般性を失なうことなく，Ω_0' を $\delta>0$ なる $(\theta_0-\delta,\theta_0+\delta)$ としてもよい．よって，$A(\theta_0,\theta)$ はこの区間上で

単調減少である．$A(\theta_0,\theta_0)=0$ より，$A(\theta_0,\theta_0-\delta)>0$ かつ $A(\theta_0,\theta_0+\delta)<0$ となる．ゆえに，$n(\delta,\varepsilon)$ が存在して θ_0 が θ の真値ならばすべての $n>n(\delta,\varepsilon)$ に対して次の両不等式が成り立つ確率は $1-\varepsilon$ を越える．

$$(12.3.3)\qquad \begin{aligned}&\frac{1}{n}S_n(x_1,\cdots,x_n\,;\,\theta)>0, &&\theta=\theta_0-\delta \text{ のとき}\\ &\frac{1}{n}S_n(x_1,\cdots,x_n\,;\,\theta)<0, &&\theta=\theta_0+\delta \text{ のとき.}\end{aligned}$$

$S(x,\theta)$ は確率 0 の集合をのぞく R_1 のすべての x に対して，$(\theta_0-\delta,\theta_0+\delta)$ 上で θ について連続だから

$$\frac{1}{n}\sum_{\xi=1}^{n}S(x_\xi\,;\,\theta),\qquad \text{すなわち}\qquad \frac{1}{n}S_n(x_1,\cdots,x_n\,;\,\theta)$$

についても同様である．ゆえに，θ_0 が θ の真値ならば，

$$(12.3.4)\qquad P\Big(\frac{1}{n}S_n(x_1,\cdots,x_n\,;\,\theta)=0,\ \text{すべての } n>n(\delta,\varepsilon) \text{ および区間 } (\theta_0\pm\delta) \text{ 内のある } \theta \text{ に対して}\,\Big|\,\theta_0\Big)>1-\varepsilon.$$

これは θ_0 に概収束する (12.3.2) の根の列が存在することと同値である．特に，ある整数 n_0 が存在して (12.3.2) が $n=n_0,n_0+1,\cdots$ に対して一意解 $\hat\theta(x_1,\cdots,x_n)$ を持てば，列 $\hat\theta(x_1,\cdots,x_n)$，$n=n_0,n_0+1,\cdots$ は θ_0 に概収束する．

まとめると

12.3.2　$(x_1,\cdots,x_n)$ を c.d.f. $F(x\,;\,\theta_0)$ からの標本とする．ただし，$F(x\,;\,\theta)$ は Ω_0 において θ の 1 次導関数に関して正則である．$S(x\,;\,\theta)$ を確率 0 の集合をのぞいた R_1 のすべての x の値に対して，θ に関して連続とする．このとき，θ_0 に概収束する (12.3.2) の解の列が存在する．特に，(12.3.2) が n_0 より大きい n に対して，一意解 $\hat\theta(x_1,\cdots,x_n)$ を持てば，列 $\hat\theta(x_1,\cdots,x_n)$，$n=n_0$，$n_0+1,\cdots$ は θ_0 に概収束する．

$S_n(x_1,\cdots,x_n\,;\,\theta)$ は尤度要素の対数の 1 次導関数 $\sum_{\xi=1}^{n}\log dF(x_\xi\,;\,\theta)$ であり，$\hat\theta(x_1,\cdots,x_n)$ は $S_n(x_1,\cdots,x_n\,;\,\theta)$ が 0 となる θ の値であることから，$\hat\theta$ が一意でかつ尤度要素を最大にするとき，θ_0 に対する**最尤推定量**と呼ばれる．この用語は Fisher (1922) により導入された．Wald (1949 *b*) は尤度を最大にする (12.3.2) の解が，ある条件下で，θ_0 の一致推定量となることを示している．最尤推定量の一致性の詳しい解析および関連する問題は Barankin と Gurland (1951)，Huzurbazar (1948)，LeCam (1956)，Wald (1948, 1949 *b*) になされている

(c) 最尤推定量の漸近分布

$F(x\,;\,\theta)$ の θ の 2 次までの導関数に関する正則性を仮定すれば，大きな n に対する $\hat\theta$ の漸近分布に関して次のように述べることができる．

12.3.3　$(x_1,\cdots,x_n)$ は c.d.f. $F(x\,;\,\theta_0)$ からの標本である．ただし $F(x\,;\,\theta)$ は Ω_0 において θ の 2 次までの導関数に関して正則である．θ_0 の最尤推定量 $\hat\theta$ が $n\geqslant$(ある n_0) に対して一意で (12.2.1) で定義した $\prod_{\xi=1}^{n}F(x_\xi\,;\,\theta)$ に関する (可測な) 確率変数ならば，$\hat\theta$ の分布は大きな n に対して漸近的に正規分布 $N(\theta_0\,;\,1/[nB^2(\theta_0,\theta_0)])$ に従う．

$S(x\,;\,\theta)$ は Ω_0 のいたるところで，0 確率の集合をのぞいた R_1 のすべての点 x に対して θ の導関数を持つので，$S_n(x_1,\cdots,x_n\,;\,\theta)$ に対しても 0 確率の集合をのぞいた R_n のすべての点 $(x_1,\cdots,x_n)$ に対して同様に θ の導関数を持つ．$n\geqslant$(ある n_0) に対して $\hat\theta$ が一意ならば，**12.3.2** より，$n\longrightarrow\infty$ のとき $\hat\theta$ は θ_0 に概収束する．さらに，$\hat\theta(x_1,\cdots,x_n)$ が確率変数ならば，任意の $\delta>0$，$\varepsilon>0$ に対して $n(\delta,\varepsilon,n_0)$ および $|\theta_0-\hat\theta(x_1,\cdots,x_n)|<\delta$ で定義される R_n の部分集合 E_n が存在して，

$$(12.3.5)\qquad P\Big((x_1,\cdots,x_n)\in E_n,\ \text{すべての } n>n(\delta,\varepsilon,n_0) \text{ に対して}\,\Big|\,\theta_0\Big)>1-\varepsilon$$

が成立する．E_n では

$$(12.3.6)\qquad \begin{aligned}&\frac{1}{\sqrt{n}}S_n(x_1,\cdots,x_n\,;\,\theta_0)\\ &\quad=\frac{1}{\sqrt{n}}S_n(x_1,\cdots,x_n\,;\,\hat\theta)+\frac{1}{\sqrt{n}}S_n'(x_1,\cdots,x_n\,;\,\theta^*)(\theta_0-\theta).\end{aligned}$$

ここで $\theta^*(x_1,\cdots,x_n)$ は

$$(12.3.7)\qquad |\theta_0-\theta^*|\leqslant|\theta_0-\theta|$$

を満たす確率変数で，

$$(12.3.8)\qquad S_n'(x_1,\cdots,x_n\,;\,\theta^*)=\sum_{\xi=1}^{n}S'(x_\xi\,;\,\theta^*)$$

かつ

$$S'(x\,;\,\theta) = \frac{\partial}{\partial\theta} S(x\,;\,\theta) \tag{12.3.9}$$

である．しかし (12.3.5), (12.3.6) の両式は

$$P\Big(\frac{1}{\sqrt{n}} S_n(x_1, \cdots, x_n\,;\,\theta_0) = \frac{1}{\sqrt{n}} S_n(x_1, \cdots, x_n\,;\,\hat\theta) + \frac{1}{n} S'_n(x_1, \cdots, x_n\,;\,\theta^*) \cdot [\sqrt{n}(\theta_0 - \hat\theta)] \ \text{すべての}\ n > n(\delta, \varepsilon, n_0)\ \text{に対して} \,\Big|\, \theta_0\Big) > 1 - \varepsilon \tag{12.3.10}$$

と同値である．定義より，$\hat\theta$ は (12.3.2) を満足するので，(12.3.10) は

$$P\Big(\frac{1}{\sqrt{n}} S_n(x_1, \cdots, x_n\,;\,\theta_0) = \frac{1}{n} S'_n(x_1, \cdots, x_n\,;\,\theta^*)[\sqrt{n}(\theta_0 - \hat\theta)] \ \text{すべての}\ n > n(\delta, \varepsilon, n_0)\ \text{に対して} \,\Big|\, \theta_0\Big) > 1 - \varepsilon \tag{12.3.11}$$

に帰着する．よって

$$\frac{1}{\sqrt{n}} S_n(x_1, \cdots, x_n\,;\,\theta_0), \qquad n = 1, 2, \cdots \tag{12.3.12}$$

および

$$\frac{1}{n} S'_n(x_1, \cdots, x_n\,;\,\theta^*)[\sqrt{n}(\theta_0 - \hat\theta)], \qquad n = 1, 2, \cdots \tag{12.3.13}$$

はどちらかが法則収束すれば両方とも法則収束する確率変数列である．さて

$$\frac{1}{n} S'_n(x_1, \cdots, x_n\,;\,\theta_0) = \frac{1}{n} \sum_{\xi=1}^{n} S'(x_\xi\,;\,\theta_0) \tag{12.3.14}$$

は

$$\int_{-\infty}^{\infty} S'(x\,;\,\theta_0)\, dF(x\,;\,\theta_0) = -B^2(\theta_0\,;\,\theta_0) \tag{12.3.15}$$

を平均として持つ母集団からの大きさ n の標本平均である．したがって **9.1.1** より，(12.3.14) の左辺は (12.3.15) に確率収束する．**4.3.7** より，$\hat\theta$ は θ_0 に確率収束するので，θ^* も同様である．ゆえに **4.3.8** を用いると，$\frac{1}{n} S'_n(x_1, \cdots, x_n\,;\,\theta^*)$ は (12.3.15) 式の右辺に確率収束する．

最後に，**4.3.6** を用いると，(12.3.12), (12.3.13) で与えられる確率変数列および列

$$\sqrt{n}(\hat\theta - \theta_0) B^2(\theta_0, \theta_0), \qquad n = 1, 2, \cdots \tag{12.3.16}$$

は，この 3 つのどれか 1 つが法則収束すればすべて法則収束する．しかし，**12.3.1** より，列 (12.3.12) は

$$N(0, B^2(\theta_0, \theta_0)) \tag{12.3.17}$$

に法則収束する．ゆえに，大きな n に対して $\hat\theta$ の漸近分布は

$$N\Big(\theta_0, \frac{1}{nB^2(\theta_0, \theta_0)}\Big). \tag{12.3.18}$$

よって **12.3.3** が得られた．(12.3.18) で与えられた $\hat\theta$ の漸近分布は Fisher (1922) により最初に求められた．

(d)　最尤推定量の漸近効率

$\tilde\theta(x_1, \cdots, x_n)$ を，(12.2.16) で定義した偏り $b_n(\theta_0)$ を持つ θ_0 に対する偏りのある推定量とする．$\tilde\theta$ が θ_0 の一致推定量ならば，この偏り $b_n(\theta_0)$ は $n \longrightarrow \infty$ のとき 0 に収束する．

n を任意に与えると，$\sigma^2(\tilde\theta_n | \theta_0)$ の下界は (12.2.17) で与えられる．よって

$$\sigma^2(\sqrt{n}\,\tilde\theta | \theta_0) = \mathscr{E}([\sqrt{n}(\tilde\theta - \theta_0 - b_n(\theta_0))]^2 | \theta_0) \geqslant \frac{[1 + b'_n(\theta_0)]^2}{B^2(\theta_0, \theta_0)}. \tag{12.3.19}$$

任意の n に対して上式の不等号が成り立つので，$n \longrightarrow \infty$ のとき $b'_n(\theta_0) \longrightarrow 0$ を仮定すれば

$$\liminf_{n\to\infty} \sigma^2(\sqrt{n}\,\tilde\theta | \theta_0) \geqslant \lim_{n\to\infty}\Big(\frac{[1 + b'_n(\theta_0)]^2}{B^2(\theta_0, \theta_0)}\Big) = \frac{1}{B^2(\theta_0, \theta_0)} \tag{12.3.20}$$

すなわち

$$\liminf_{n\to\infty} \sigma^2(\sqrt{n}\,\tilde\theta | \theta_0) \geqslant \frac{1}{B^2(\theta_0, \theta_0)}. \tag{12.3.21}$$

これにより，次の結果が得られる．

12.3.5　$(x_1, \cdots, x_n)$ を c.d.f. $F(x\,;\,\theta_0)$ からの標本とする．ただし，$F(x\,;\,\theta)$ は θ の 1 次導関数に関して正則で，Ω_0 の θ に対して $B^2(\theta, \theta)$ が存在するとする．$\tilde\theta$ が θ_0 に対する一致推定量であり，$\tilde\theta$ の偏りの θ に関する導関数が $n \longrightarrow \infty$ のとき極限 0 を持つならば，$\sigma^2(\sqrt{n}\,\tilde\theta)$ の $n \longrightarrow \infty$ のときの極限は，$1/B^2(\theta_0, \theta_0)$ 未満にはならない．

θ に対する一致推定量の重要な場合として **12.3.3** の条件下における最尤推定量 $\hat\theta$ がある．また **12.2.3** より，

$$\lim_{n\to\infty} \sigma^2(\sqrt{n}\,\hat\theta | \theta_0) = \frac{1}{B^2(\theta_0, \theta_0)} \tag{12.3.22}$$

になる．

$n \longrightarrow \infty$ のときの $\tilde{\theta}$ の極限効率あるいは漸近効率は

(12.3.23)
$$\mathrm{Ieff}\,(\tilde{\theta}|\theta_0) = \frac{B^{*2}(\theta_0, \theta_0)}{B^2(\theta_0, \theta_0)}$$

で定義される.

よって，明らかに

(12.3.24)
$$\mathrm{Ieff}\,(\hat{\theta}|\theta_0) = 1.$$

$\mathrm{Ieff}\,(\tilde{\theta}|\theta_0) = 1$ なる θ_0 の推定量 $\tilde{\theta}$ は，θ_0 に対する漸近的有効推定量という．$\hat{\theta}$ についてはそうである．よって

12.3.6 **12.3.3** の仮定の下では，θ_0 に対する最尤推定量 $\hat{\theta}$ は一致推定量であり，かつ漸近的に有効である．

(12.2.15) で定義した効率と，(12.3.23) で定義した漸近効率の違いに注意しよう．前者の定義は任意の n に対して成り立つが，有効推定量が存在するときのみ定義される．後者は $n \longrightarrow \infty$ のときの極限として定義される性質で，漸近的有効推定量 $\hat{\theta}$ が存在するとき，すなわち **12.3.3** の仮定の下で定義される．また，一致推定量 $\tilde{\theta}$ の漸近効率は $\tilde{\theta}$ に含まれている θ_0 に関する情報量のフィッシャーの概念の大標本の場合への拡張とみなされる．

最尤推定量の漸近的性質に関する詳細な研究は，Bahadur (1958), Barankin と Gurland (1951), Kraft と LeCam (1956), Rao (1949), Wald (1948) 等によりなされている．

例題 推定量の漸近効率の例を示そう．$(x_1, \cdots, x_n)$ を正規分布 $N(\mu, \sigma^2)$ からの標本とする．$\tilde{\mu}$ を標本メディアン，$\hat{\mu}$ を標本平均とする．**9.6.6** から $p = \frac{1}{2}$ に対して，大きな n に対する $\tilde{\mu}$ の漸近分布は $N(\mu, \pi\sigma^2/2n)$ である．μ の有効推定量 $\hat{\mu}$ が n の任意の値に対して存在し，$\hat{\mu}$ は標本平均であることは明らかである．さらに **8.3.3 d** から，標本平均の分布は正規分布 $N(\mu, \sigma^2/n)$ になることがわかっている．よって，(12.3.23) を用いると，μ の真値 μ_0 を推定する標本メディアン $\tilde{\mu}$ の漸近効率は

$$\mathrm{Ieff}\,(\tilde{\mu}|\mu_0) = \frac{\sigma^2}{(\pi/2)\sigma^2} = 0.637$$

で与えられる．これは μ, σ^2 の真値に関係なく，大きな n に対して，大きさ $(2/\pi)n$ の標本平均は大きさ n の標本メディアンと同程度正確に正規分布 $N(\mu, \sigma^2)$ の平均 μ の真値を推定することを意味している．

(e) 漸近的分散 $\frac{1}{n}$ を持つ，最尤推定量の関数

12.3.3 より，$F(x\,;\,\theta)$ の正則性の条件下では，最尤推定値 $\hat{\theta}$ は大きな n に対して，漸近的分散として $1/[nB^2(\theta_0, \theta_0)]$ を持つ．ここで θ_0 は θ の真値である．理論上および実用上重要な問題は，$\zeta(\hat{\theta})$ の漸近的分散が $\frac{1}{n}$ になるような θ のある関数（これを $\zeta(\theta)$ と表わす）を新しいパラメータ ζ として用いることができるかどうか，である．

この種の関数はある条件下では求められる．$\zeta(\theta)$ を求めるために次のようにすればよい．$\zeta(\theta)$ を Ω_0 において導関数 $\zeta'(\theta)$ および唯一の逆関数 $\theta(\zeta)$ を持つ関数とする．さらに $\theta(\zeta)$ は ζ に関する導関数 $\theta'(\zeta)$ を持つこととする．ζ_0 を ζ の真値，すなわち $\theta(\zeta_0) = \theta_0$ とする．(12.1.8) を用いると，

(12.3.25)
$$B^2(\theta, \theta) = \int_{-\infty}^{\infty} (S(x\,;\,\theta))^2\, dF(x\,;\,\theta)$$

と書ける．また

(12.3.26)
$$S(x\,;\,\theta) = \frac{d}{d\theta} \log dF(x\,;\,\theta)$$

より

(12.3.27)
$$S(x\,;\,\theta(\zeta)) = \frac{d}{d\zeta} \log dF(x\,;\,\theta(\zeta)) \cdot \zeta'(\theta)$$

と表わされる．これを (12.3.25) に代入すると

(12.3.28)
$$B^2(\theta(\zeta), \theta(\zeta)) = [\zeta'(\theta)]^2 \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial\zeta} \log dF(x\,;\,\theta(\zeta))\right]^2 dF(x\,;\,\theta(\zeta)).$$

大きな n に対して漸近的分散 $\frac{1}{n}$ を持つような最尤推定量 $\hat{\zeta}$ (ζ_0 に対して) を求めることは，(12.3.28) 式の右辺の積分値が 1 になることと同値である．よって関数 $\zeta(\theta)$ は微分方程式

(12.3.29)
$$\left(\frac{d\theta}{d\zeta}\right)^2 = \frac{1}{B^2(\theta(\zeta), \theta(\zeta))}$$

すなわち

(12.3.30)
$$\zeta(\theta) = \int_{\theta_0}^{\theta} B(\theta, \theta)\, d\theta + \zeta_0, \qquad B(\theta, \theta) = +\sqrt{B^2(\theta, \theta)}$$

を満足しなければならない．まとめると次の結果が得られる．

12.3.7 $(x_1, \cdots, x_n)$ が c.d.f. $F(x\,;\,\theta_0)$ からの標本で，$\hat{\theta}$ が θ_0 に対する最尤推定量ならば，$\zeta(\hat{\theta})$ は大きな n に対して，漸近的に正規分布 $N\left(\zeta_0, \frac{1}{n}\right)$ に従って分布する．ここで $F(x\,;\,\theta)$ は θ の 2 次導関数に関して正則とする．ただ

し $\zeta(\theta)$ は (12.3.30) で与えられ，Ω_0 において導関数 $\zeta'(\theta)$ と唯一の逆関数 $\theta(\zeta)$ を持つとする．

例題　$(x_1, \cdots, x_n)$ を p.f. が

$$p(x\,;\mu_0) = \frac{\mu_0^x e^{-\mu_0}}{x\,!}, \qquad x = 0, 1, \cdots$$

なるポアソン分布 $Po(\mu_0)$ からの標本とする．このとき $\zeta(\hat{\mu})$ が $n \longrightarrow \infty$ のとき漸近的に正規分布 $N(\zeta(\mu_0), 1/n)$ を持つような μ の関数 $\zeta(\mu)$ を求めたい．12.2 (c) 節の例題から，μ_0 に対する最尤推定量 $\hat{\mu}$ は標本平均 $\bar{x}$ であることがわかっている．(12.3.25) をポアソン分布（すなわち $dF(x\,;\mu) = p(x\,;\mu)$）の場合に用いると

$$B^2(\mu, \mu) = \sum_{x=0}^{\infty} \left(\frac{x}{\mu} - 1\right)^2 p(x, \mu) = \frac{1}{\mu}$$

となる．$\zeta_0 = 2\sqrt{\mu_0}$ として (12.3.30) を適用すると，

$$\zeta(\mu) = 2\sqrt{\mu}$$

となる．ゆえに **12.3.7** より，$(x_1, \cdots, x_n)$ がポアソン分布 $Po(\mu_0)$ を持つ母集団からの標本のとき，$2\sqrt{\bar{x}}$ は大きな n に対する漸近分布として正規分布 $N(2\sqrt{\mu_0}, 1/n)$ を持つ．

12.4 区間推定量

(a) 信頼区間の定義

$(x_1, \cdots, x_n)$ を c.d.f. $F_n(x_1, \cdots, x_n\,;\theta)$ を持つ確率変数としよう．$\underline{\theta}(x_1, \cdots, x_n)$, $\bar{\theta}(x_1, \cdots, x_n)$ を $\underline{\theta} < \bar{\theta}$ なる標本要素の関数（確率変数）とする．$\underline{\theta}$, $\bar{\theta}$ を与えられた γ に対して

(12.4.1) $$P(\underline{\theta} < \theta < \bar{\theta} \,|\, \theta_0) = \gamma$$

となるように選ぶことができれば（ただし $P(\underline{\theta} < \theta < \bar{\theta} \,|\, \theta)$ は $F_n(x_1, \cdots, x_n\,;\theta)$ から定まる確率を表わす），$(\underline{\theta}, \bar{\theta})$ を θ に対する $100\gamma\%$ **信頼区間**という．ここで $\underline{\theta}$, $\bar{\theta}$ をそれぞれ θ に対する**下側，上側信頼限界**といい，γ を**信頼係数**という．$(\underline{\theta}, \bar{\theta})$ はこの区間が θ の真値を含む $F_n(x_1, \cdots, x_n\,;\theta)$ による確率が γ になる 2 次元確率変数である．正規分布からの標本抽出に対する信頼区間の例は 10.2 (c), 10.4, 10.5, 10.6 の各節にすで

に与えられている．信頼区間の数学的定式化は Laplace (1814) によって最初に導入された．そこでは 2 項分布 (6.2.2) に従う確率変数 x の観測値から分布の p の値を推定する問題を扱っている．ラプラスは信頼区間を固定し，p を確率変数とみなした．のちに，この方法は Wilson (1927) によって再度検討され，そこでは確率区間としての区間の正しい解釈を与えている．近代理論への展開および信頼区間という用語は Neyman (1937) により導入された．

(b) 連続な c.d.f. からの標本による信頼区間の構成方法

連続な c.d.f. $F(x\,;\theta)$ から抽出された標本から信頼区間を構成する方法は次のように述べられる．

12.4.1　$(x_1, \cdots, x_n)$ を連続な c.d.f. $F(x\,;\theta)$ からの標本とする．$g(x_1, \cdots, x_n\,;\theta)$ は次の (i)〜(iii) を満足するとする．(i) θ_0 を含む区間 (θ_1, θ_2) の各点 θ と，確率 0 の集合をのぞいた標本空間 R_n の各点で定義されている．(ii) θ に関して連続で，かつ θ に関して単調増加あるいは単調減少である．(iii) θ に依存しない c.d.f. を持つ．また (g_1, g_2) を $P(g_1 < g < g_2 \,|\, \theta) = \gamma$ なる区間とする．このとき，θ_0 が θ の真値ならば，等式 $g(x_1, \cdots, x_n\,;\theta) = g_1, g_2$ の解 $\underline{\theta}$, $\bar{\theta}$（ただし $\underline{\theta} < \bar{\theta}$）が存在して，$(\underline{\theta}, \bar{\theta})$ は θ_0 に対する $100\gamma\%$ 信頼区間になる．

12.4.1 を確かめよう．まず $g(x_1, \cdots, x_n\,;\theta)$ の c.d.f. は θ に依存しないから，与えられた $0 < \gamma < 1$ と与えられた真値 θ_0 に対して，θ_0 が θ の真値のとき

(12.4.2) $$P(g_1 < g(x_1, \cdots, x_n\,;\theta_0) < g_2 \,|\, \theta_0) = \gamma$$

となるように，g_1, g_2 を選ぶことができる（もちろん，多くの方法がある）．(θ_1, θ_2) 上の θ で，$g(x_1, \cdots, x_n\,;\theta)$ は連続かつ単調増加（または減少）なので，(12.4.2) は明らかに

(12.4.3) $$P(\underline{\theta} < \theta_0 < \bar{\theta} \,|\, \theta_0) = \gamma$$

と同値である．ここに $\underline{\theta} < \bar{\theta}$ は方程式 $g(x_1, \cdots, x_n\,;\theta) = g_1, g_2$ の解で，もちろん確率変数である．

次は $F(x\,;\theta)$ の x に関する連続性を仮定して，c.d.f. が θ に無関係になるような関数 $g(x_1, \cdots, x_n\,;\theta)$ を求める問題である．このような関数はつねに求められる．事実，単純無作為標本に対する $F_n(x_1, \cdots, x_n\,;\theta)$，すなわち，$\prod_{\xi=1}^{n} F(x_\xi\,;\theta)$ はこのような関数に

なる．ただし $F(x\,;\theta)$ は確率 0 の集合をのぞくすべての点 x で θ に関して連続でかつ単調増加または減少であるとする．なぜなら，確率変数 $F(x_\xi\,;\theta)$, $\xi=1,2,\cdots,n$ は独立で，8.7 (a) 節で指摘したようにおのおの，矩形分布 $R\left(\frac{1}{2},1\right)$ に従って分布するからである．さらに，$-\log F(x_\xi\,;\theta)$ はガンマ分布 $G(1)$ を持ち，ガンマ分布の再生性より，$-\sum_{\xi=1}^{n}\log F(x_\xi\,;\theta)$ はガンマ分布 $G(n)$ を持つ．それゆえ

(12.4.4)
$$P\left(-\log b_2<-\sum_{\xi=1}^{n}\log F(x_\xi\,;\theta)<-\log b_1|\theta\right)$$
$$=\frac{1}{\Gamma(n)}\int_{-\log b_2}^{-\log b_1}y^{n-1}e^{-y}\,dy.$$

b_1 と b_2 を適当に選べば，右辺の積分は γ になる．よって

(12.4.5)
$$P\left(b_1<\prod_{\xi=1}^{n}F(x_\xi\,;\theta)<b_2|\theta\right)=\gamma.$$

さて，R_1 のすべての点 x（確率 0 の集合をのぞく）に対して $F(x\,;\theta)$ が θ に関して連続で単調増加（または減少）ならば，$\prod_{\xi=1}^{n}F(x_\xi\,;\theta)$ についても同様である．それゆえ (12.4.5) の（　）内の不等式は

(12.4.6)
$$P(\underline{\theta}<\theta_0<\bar{\theta}|\theta_0)=\gamma$$

と書ける．これで θ_0 に対する $100\gamma\%$ 信頼区間が求められた．

12.4.1 で $g(x_1,\cdots,x_n\,;\theta)$ の単調性の仮定をはずしても，(12.4.2) は成り立つ．しかし，(12.4.2) の不等号が逆向きになれば，信頼区間 $(\underline{\theta},\bar{\theta})$ のかわりに (θ_1,θ_2) 内の確率集合 E になる．ただし E は $(x_1,\cdots,x_n)$ に依存するが，区間 $(\underline{\theta},\bar{\theta})$ には依存しない．よって (12.4.3) は $P(\theta\in E|\theta)=\gamma$ で置き換えられる．ランダムであるのは θ でなくて E であることに十分注意しておこう．θ_0 の推定量という観点からすれば，確率区間 $(\underline{\theta},\bar{\theta})$ は確率集合 E よりも有用である．θ_0 に対する区間推定の概念の集合推定への拡張は 12.7 節で区間推定量を多次元パラメータ θ に拡張する問題を考えるとき有効になってくる．

(c)　離散型分布からの標本による信頼区間

$F(x\,;\theta)$ が離散型確率変数の c.d.f. である場合には，もちろん **12.4.1** は適用できない．すなわち任意の γ に対して，γ にちょうど等しい信頼係数を持つ信頼区間を求めることができない．しかし，ある条件下では，信頼係数が γ より少さくならない信頼区間は求められる．もっとも，この場合は連続型確率変数の場合ほどエレガントではない．

12.4.2　$(x_1,\cdots,x_n)$ を c.d.f. $F(x\,;\theta)$ からの標本とする．ただし x は離散型確率変数で θ はパラメータ空間 (θ_1,θ_2) のパラメータである．$\tilde{\theta}$ を標本空間 R_n のすべての点で定義された (θ_1,θ_2) 内に値を持つ θ の推定量とする．また $V(\tilde{\theta}\,;\theta)$ を $\tilde{\theta}$ の c.d.f. とし，$V^*(\tilde{\theta}\,;\theta)=1-V(\tilde{\theta}\,;\theta)$ とする．さらに，$V(\tilde{\theta}\,;\theta)$ は $\tilde{\theta}$ の各質点で，θ に関して連続で減少関数であるとする．すなわちすべての $\tilde{\theta}\in(\theta_1,\theta_2)$ に対して $\lim_{\theta\to\theta_1}V(\tilde{\theta}\,;\theta)=1$, $\lim_{\theta\to\theta_2}V(\tilde{\theta}\,;\theta)=0$ である．さらに $\underline{\theta}$, $\bar{\theta}$ をそれぞれ $V(\tilde{\theta}\,;\theta)=\gamma_1$, $V^*(\tilde{\theta}\,;\theta)=\gamma_1^*$ に対応する θ の値とする．ただし γ_1, γ_1^* は非負で，$0<\gamma=1-\gamma_1-\gamma_1^*<1$ とする．このとき，$(\underline{\theta},\bar{\theta})$ は信頼係数が $\geqslant\gamma$ なる θ_0 の信頼区間である．

12.4.2 を確かめよう．まず θ_1 を $V(\tilde{\theta}\,;\theta_0)\leqslant\gamma_1$ なる $\tilde{\theta}$ の最大値とし，θ_2 を $V^*(\tilde{\theta}\,;\theta_0)\leqslant\gamma_1^*$ なる $\tilde{\theta}$ の最小値とする．このとき

(12.4.7)
$$P(\theta_1<\tilde{\theta}<\theta_2)\geqslant\gamma.$$

ここで $\gamma=1-\gamma_1-\gamma_1^*$．しかし $V(\tilde{\theta}\,;\theta)$ は θ に関して単調減少で $\tilde{\theta}$ に関しては非減少であり，$V^*(\tilde{\theta}\,;\theta)$ は θ に関して単調増加で $\tilde{\theta}$ に関しては非増加であるので，$V(\tilde{\theta}\,;\theta_0)>\gamma_1$, $V^*(\tilde{\theta}\,;\theta_0)>\gamma_1^*$ のとき，すなわち $\underline{\theta}\leqslant\theta_0\leqslant\bar{\theta}$ のときかつそのときのみ $\theta_1<\tilde{\theta}<\theta_2$ となることは明らかである．ゆえに

(12.4.8)
$$P(\underline{\theta}\leqslant\theta_0\leqslant\bar{\theta}|\theta_0)\geqslant\gamma.$$

よって，$(\underline{\theta},\bar{\theta})$ は信頼係数 $\geqslant\gamma$ なる θ の信頼区間である．

例題　2 項分布 $Bi(1,p_0)$ を持つ母集団からの標本 $(x_1,\cdots,x_n)$ を考えよう．この分布の p.f. は

$$f(x\,;p_0)=p_0^x(1-p_0)^{1-x}$$

で，質点は $x=0,1$, p_0 は $(0,1)$ 上の数である．標本平均 $\bar{x}$ は標本空間 $0,1/n,\cdots,n/n$ 上で 2 項分布 $Bi(n,p_0)$ を持ち，p_0 の一致推定量である．さらに標本抽出される母集団が 2 項分布 $Bi(1,p)$ を持つときには，$\bar{x}$ の c.d.f. は

$$V(\bar{x}\,;p)=\sum_{i=0}^{n\bar{x}}\binom{n}{i}p^i(1-p)^{n-i}$$

で定義され，$V^*(\bar{x}\,;p)$ は

$$V^*(\bar{x}\,;p)=1-V(\bar{x}\,;p)$$

で定義される．

さて，$0 < n\bar{x} < p$ に対して，$V(\bar{x}\,;\,p)$ は p に関して単調減少であり，$V^*(\bar{x}\,;\,p)$ は p に関して単調増加である．なぜなら $(d/dp)V(\bar{x}\,;\,p) < 0$，$(d/dp)V^*(\bar{x}\,;\,p) > 0$ であるからである．さらに，すべての $\bar{x} \in (0,1)$ に対して $V(\bar{x}\,;\,0) = 1$ かつ $V(\bar{x}\,;\,1) = 0$ である．ゆえに **12.4.2** を適用すると，$\underline{p}$ と $\overline{p}$ がそれぞれ $V(\bar{x}\,;\,p_0) = \gamma_1$ と $V^*(\bar{x}\,;\,p_0) = \gamma_1^*$ の解で，p_0 が母集団の p の真値で，$\gamma = 1 - \gamma_1 - \gamma_1^*$ ならば

$$P(\underline{p} \leqslant p_0 \leqslant \overline{p} \,|\, p_0) \geqslant \gamma$$

は明らかである．すなわち $(\underline{p}, \overline{p})$ は信頼係数 $\geqslant \gamma$ なる p_0 の信頼区間である．

上で求めた p に対するこの形の信頼区間は Clopper と Peason (1934) が $\gamma_1 = \gamma_1^* = 0.05, 0.025$ で，n を 10 から 1000 までの値について図示している．

上の例題と同じ方法を用いて，Garwood (1936), Ricker (1937) がポアソン分布におけるパラメータ μ の信頼区間を求めている．

12.4.2 の方法はパラメータ空間が区間 (θ_1, θ_2) の離散部分集合，およびこの区間のある特定の部分集合の場合にも拡張できる．たとえばこのような例としては確率変数 x が 6.1 (a) 節で定義された超幾何分布 $H(N, n\,;\,p)$（ただし N と n は既知）を持つ場合に p の信頼区間を求める場合がある．この場合 p のパラメータ空間 Ω は離散的な点 0, $1/N, 2/N, \cdots, 1$ からなっている．この場合における信頼区間は，Chung と DeLury (1950) が $N = 500, 2500, 10000$，$n/N = 0.05, 0.10\,(0.10)\,0.90$，$\gamma = 0.90, 0.95, 0.99$ の場合について計算している．

(d)　フィドーシャル確率とフィドーシャル区間に関する注意

Fisher (1935 *b*) が導入した概念で，信頼区間に関してフィドーシャル区間という概念がある．θ のパラメータ空間を区間 (θ_1, θ_2) と仮定し，s を標本空間が (a, b) なる統計量とする．$v(s\,;\,\theta)$ を s の p.d.f. とし，$V(s\,;\,\theta)$ を c.d.f. とする．$V^*(\theta\,;\,s) = 1 - V(s\,;\,\theta)$ が (a, b) 内の各 s に対し，θ に関して連続で狭義単調増加，したがって (a, b) 内の各 s に対して $\lim_{\theta \to \theta_1} V^*(\theta\,;\,s) = 0$ かつ $\lim_{\theta \to \theta_2} V^*(\theta\,;\,s) = 1$ であれば，(θ_1, θ_2) 上の θ の関数として $V^*(\theta\,;\,s)$ は (a, b) 内の各 s に対して c.d.f. としてのすべての形式的な性質を満足する．これを統計量 s に基づく θ のフィドーシャル c.d.f. という．したがって，θ', θ'' $(\theta' < \theta'')$ を $V^*(\theta'\,;\,s) = \gamma_1$，$V^*(\theta''\,;\,s) = \gamma_2$ となるように選べば（ただし $\gamma_2 - \gamma_1 = \gamma$，$0 < \gamma < 1$ とする），フィドーシャル確率式

$$\text{fid } P(\theta' < \theta < \theta'' \,|\, s) = \gamma \tag{12.4.9}$$

が得られる．

(θ', θ'') は θ に対する $100\gamma\,\%$ フィドーシャル区間と呼ばれる．θ_0 が θ の真値で，s'，s'' をそれぞれ等式 $V(s\,;\,\theta_0) = 1 - \gamma_2$ と $V(s\,;\,\theta_0) = 1 - \gamma_1$ を満足する s の値とする．ただし $\gamma_2 - \gamma_1 = \gamma$．このとき，$P(s' < s < s'' \,|\, \theta_0) = \gamma$ となる．しかし s が (s', s'') 内にあるときかつそのときにのみ θ_0 は (θ', θ'') 内にある．よって (θ', θ'') はまた θ_0 の $100\gamma\,\%$ 信頼区間である．

フィドーシャル分布およびフィドーシャル区間を生成する別の方法としてはピボット関数によるものがある．すなわち，統計量 s と上述のパラメータに対して，$g(s\,;\,\theta)$ が各 s に対して θ の狭義単調増加関数で，各 θ に対しても s の関数として同じ性質を持っているとし，$g(s\,;\,\theta)$ は高々 g に含まれる θ にだけしか依存しない確率要素 $h(g)\,dg$ を持つとする．$g(s\,;\,\theta)$ が s と θ に関して 1 次偏導関数を持てば，$h(g)(\partial g/\partial s)\,ds$ は s の確率要素となり，$h(g)(\partial g/\partial\theta)\,d\theta$ はパラメータ θ のフィドーシャル確率要素である．ここで g_1 と g_2 $(g_1 < g_2)$ を

$$P(g_1 < g(s\,;\,\theta) < g_2 \,|\, \theta) = \gamma \tag{12.4.10}$$

なる数としよう．s_1 と s_2 $(s_1 < s_2)$ をそれぞれ $g(s\,;\,\theta) = g_1, g_2$ を満足する s の値とする．一方，θ_1, θ_2 $(\theta_1 < \theta_2)$ をそれぞれ $g(s\,;\,\theta) = g_1, g_2$ を満たす θ の値とする．このとき $\theta \in (\theta_1, \theta_2)$ となるための必要十分条件は $s \in (s_1, s_2)$ である．ゆえに

$$\int_{\theta_1}^{\theta_2} h(g)\frac{\partial g}{\partial\theta}\,d\theta = \int_{s_1}^{s_2} h(g)\frac{\partial g}{\partial s}\cdot ds$$

となる．ここに左辺はフィドーシャル確率 fid $P(\theta_1 < \theta < \theta_2)$ であり，右辺は確率 $P(s_1 < s < s_2)$ である．これと同じ方法が，$g(s\,;\,\theta)$ が s に関して狭義単調増加で，θ に関して減少の場合，またはこの反対の場合，あるいは $g(s\,;\,\theta)$ が両方の変数に対して減少の場合にも適用できる．

注意　上で示したフィドーシャル分布，フィドーシャル区間，信頼区間の構成に関する前述の 2 つの方法は同等である．しかし，フィドーシャル区間が信頼区間に同値でない場合がある．これは，フィドーシャル区間をどう解釈するかという問題提起になる．この問題の核心は 2 つの分布からのそれぞれの標本の標本平均と標本分散を含んでいるような統計量，たとえば 2 つの正規分布 $N(\mu_1, \sigma_1^2)$，$N(\mu_2, \sigma_2^2)$ の平均の差 $\mu_1 - \mu_2$ のフィドーシャル分布の意味にある．これを例について正確に表現しよう．n_1，$\bar{x}_1$，s_1^2 と n_2，$\bar{x}_2$，s_2^2 をそれぞれ $N(\mu_1, \sigma_1^2)$，$N(\mu_2, \sigma_2^2)$ からの独立な標本の大きさ，平均，分散とする．このとき

(12.4.11)　　$t_1 = \sqrt{n_1}(\bar{x}_1 - \mu_1)/s_1, \qquad t_2 = \sqrt{n_2}(\bar{x}_2 - \mu_2)/s_2$

はそれぞれ自由度 $n_1 - 1$ と $n_2 - 1$ を持つ独立なスチューデント比である．$f_m(t)\,dt$ を自由度 m のスチューデント比 t の確率要素とすれば，t_1 と t_2 の確率要素は $f_{n_1-1}(t_1)f_{n_2-1}(t_2)\,dt_1\,dt_2$ である．(12.4.11) の t_1, t_2 で表わされる量をピボット量として用いれば，与えられた n_1, $\bar{x}_1$, s_1^2, n_2, $\bar{x}_2$, s_2^2 に対する (μ_1, μ_2) のフィドーシャル確率要素

(12.4.12)　　$f_{n_1-1}(\sqrt{n_1}(\bar{x}_1 - \mu_1)/s_1)f_{n_2-1}(\sqrt{n_2}(\bar{x}_2 - \mu_2)/s_2)[\sqrt{n_1 n_2}/s_1 s_2]\,d\mu_1\,d\mu_2$

が得られる．(12.4.12) から $(\mu_1 - \mu_2)$ のフィドーシャル確率分布が求まり，したがってこの分布から $(\mu_1 - \mu_2)$ に対するフィドーシャル区間が求まる．しかし，ある種の正則条件を満たす関数 $g(\bar{x}_1, \bar{x}_2, s_1^2, s_2^2, \mu_1 - \mu_2)$ の中で分布が n_1, n_2, g にのみ依存し，この分布から $\mu_1 - \mu_2$ に対する信頼区間を求めうるような関数 g は存在しないことが示せる．よってこの例題のフィドーシャル区間は信頼区間と異なっていることがわかる．問題はこの場合のフィドーシャル区間をどう解釈するかである．

これはベールンス＝フィッシャー問題［Fisher (1935 *b*) を参照］として知られている．この問題に興味を持たれる読者は Tukey (1957 *c*) およびこの論文の参考文献を参照されたい．

12.5　大標本による区間推定

(a)　尤度推定関数による区間推定

大標本による母集団パラメータの区間推定に関する漸近論は，12.3 節で示したように，大標本による点推定論をそのまま適用して得られる．漸近的有効点推定はあとでわかるように，大標本の場合には漸近的最短区間推定となる．

$(x_1, \cdots, x_n)$ を c.d.f. $F(x\,;\,\theta_0)$ を持つ分布からの標本とする．便宜上，尤度推定関数 $h_n(x_1, \cdots, x_n\,;\,\theta)$ を

(12.5.1)　　$$h_n(x_1, \cdots, x_n\,;\,\theta) = \frac{S_n(x_1, \cdots, x_n\,;\,\theta)}{nB(\theta_0, \theta)}$$

で定義しよう．ここに $B(\theta_0, \theta) = +\sqrt{B^2(\theta_0, \theta)}$. **12.3.1** の仮定の下では，$\theta$ の真値が θ_0 のとき，$\sqrt{n}\,h_n(x_1, \cdots, x_n\,;\,\theta_0)$, $n \longrightarrow \infty$ は極限分布として分布 $N(0,1)$ を持つ．ここで **12.3.3** の強い条件を採用すると，すべての $n > n_\varepsilon$ に対して E_n ((12.3.5) で定義されている) の点 $(x_1, \cdots, x_n)$ で

$$\sqrt{n}\,h_n(x_1, \cdots, x_n\,;\,\theta_0) = h_n'(x_1, \cdots, x_n\,;\,\theta^*)[\sqrt{n}(\theta_0 - \hat{\theta})]$$

となる確率は $1 - \varepsilon$ を越えない．ただし $h_n'(x_1, \cdots, x_n\,;\,\theta) = (\partial/\partial\theta)h_n(x_1, \cdots, x_n\,;\,\theta)$ で，$\hat{\theta}$ と θ^* は (12.3.7) で定義されている．したがって，2 つの確率変数列

(12.5.2)　　$\sqrt{n}\,h_n(x_1, \cdots, x_n\,;\,\theta_0), \qquad n = 1, 2, \cdots$

(12.5.3)　　$h_n'(x_1, \cdots, x_n\,;\,\theta^*)[\sqrt{n}(\theta_0 - \hat{\theta})], \qquad n = 1, 2, \cdots$

はともに正規分布 $N(0,1)$ に法則収束する．さらに $(\partial/\partial\theta)B^2(\theta_0, \theta)$ が存在し，Ω_0 上で有界であって，しかも $\theta = \theta_0$ ならば，

(12.5.4)　　$h_n'(x_1, \cdots, x_n\,;\,\theta^*), \qquad h_n'(x_1, \cdots, x_n\,;\,\hat{\theta})$

は $n \longrightarrow \infty$ のとき $B(\theta_0, \theta_0)$ に確率収束する．それゆえ

$$\lim_{n\to\infty} P[-\lambda_\gamma < h_n'(x_1, \cdots, x_n\,;\,\hat{\theta})[\sqrt{n}(\theta_0 - \hat{\theta})] < +\lambda_\gamma|\theta_0] = \gamma$$

と書ける．ここで $\Phi(\lambda_\gamma) - \Phi(-\lambda_\gamma) = \gamma$. この結果

(12.5.5)
$$\lim_{n\to\infty} P\left[\hat{\theta} - \frac{\lambda_\gamma}{\sqrt{n}\,h_n'(x_1, \cdots, x_n\,;\,\hat{\theta})} < \theta_0 < \hat{\theta} + \frac{\lambda_\gamma}{\sqrt{n}\,h_n'(x_1, \cdots, x_n\,;\,\hat{\theta})}\middle|\theta_0\right] = \gamma$$

となる．端点 $\hat{\theta} \pm \dfrac{\lambda_\gamma}{\sqrt{n}\,h_n'(x_1, \cdots, x_n\,;\,\hat{\theta})}$ を持つ区間を大きな n に対する θ_0 の漸近的 $100\gamma\,\%$ 信頼区間という．その長さは $\dfrac{2\lambda_\gamma}{\sqrt{n}\,h_n'(x_1, \cdots, x_n\,;\,\hat{\theta})}$ である．

(b)　一般の推定関数による区間推定

この区間の長さの平方と，ある種の条件を満たす任意の推定関数 $g_n(x_1, \cdots, x_n\,;\,\theta)$ を用いて得られた区間の長さの平方との比は 1 を越えない数に確率収束することを見よう．

$(x_1, \cdots, x_n)$ をその最初の 2 つの θ の導関数に関して正則な $F(x\,;\,\theta)$ からの標本とする．F_n を $\prod_{\xi=1}^{n} F(x_\xi\,;\,\theta)$ とし，F_{n0} を $\theta = \theta_0$ における F_n の値とする．

(12.5.6)　　$g_n(x_1, \cdots, x_n\,;\,\theta)$

を次の (i) から (v) を満たす確率変数とする．

(12.5.7)　(i)　$\int_{R_n} [\sqrt{n}\,g_n(x_1, \cdots, x_n\,;\,\theta)]^j\,dF_n = 0, 1\,;\ j = 1, 2,$ Ω_0 の θ に対して

(ii)　$\sqrt{n}\,g_n(x_1, \cdots, x_n\,;\,\theta_0)$ は θ_0 が θ の真値ならば，$n \longrightarrow \infty$ のとき極限分布として $N(0,1)$ を持つ．

(iii)　$g_n(x_1, \cdots, x_n\,;\,\theta)$ は Ω_0 における θ と標本空間 R_n における 0 確率

の集合をのぞいたすべての点に対して連続な θ の導関数 $g_n'(x_1, \cdots, x_n\,;\,\theta)$ を持つ．

(iv)　θ_0 が θ の真値ならば，$g_n'(x_1, \cdots, x_n\,;\,\theta)$ は Ω_0 における θ に関して一様に $B^*(\theta_0, \theta)$ に確率収束する．ただし $B^*(\theta_0, \theta)$ は Ω_0 上で有界で，$B^*(\theta_0, \theta_0) \neq 0$ である．

(v)
$$A_n^*(\theta, \theta') = \int_{R_n} g_n(x_1, \cdots, x_n\,;\,\theta')\,dF_n$$
$$B_n^*(\theta, \theta') = \int_{R_n} g_n'(x_1, \cdots, x_n\,;\,\theta')\,dF_n$$

ならば

(12.5.8)
$$\begin{cases} \dfrac{\partial}{\partial\theta'} A_n^*(\theta, \theta') = B_n^*(\theta, \theta'), & \Omega_0 \times \Omega_0 \text{ の点 } (\theta, \theta') \text{ に対して} \\ A_n^*(\theta, \theta) = 0, & \Omega_0 \text{ の点 } \theta \text{ に対して} \\ \lim\limits_{n\to\infty} B_n^*(\theta, \theta) = B^*(\theta, \theta), & \Omega_0 \text{ の点 } \theta \text{ に対して.} \end{cases}$$

上に示した性質を持つ関数 $g_n(x_1, \cdots, x_n\,;\,\theta)$ を θ_0 に対する**正則推定関数**という．(12.5.1) で定義された関数 $h_n(x_1, \cdots, x_n\,;\,\theta)$ は θ_0 の正則推定関数である．

12.3.2 と同様に議論すれば，$n \geqq$ (ある n_0) に対して方程式 $g_n(x_1, \cdots, x_n\,;\,\theta) = 0$ が一意解を持てば，方程式の列

(12.5.9)
$$g_n(x_1, \cdots, x_n\,;\,\theta) = 0, \qquad n = n_0, n_0+1, \cdots$$

は θ_0 に確率収束する解の列

(12.5.10)
$$\tilde\theta(x_1, \cdots, x_n), \qquad n = n_0, n_0+1, \cdots$$

を持つことが証明できる．さらに 2 つの確率変数列

(12.5.11)
$$\sqrt{n}\, g_n(x_1, \cdots, x_n\,;\,\theta_0), \qquad n = n_0, n_0+1, \cdots$$

(12.5.12)
$$g_n'(x_1, \cdots, x_n\,;\,\tilde\theta^*)[\sqrt{n}(\theta_0 - \tilde\theta)], \qquad n = n_0, n_{0+1}, \cdots$$

はともに正規分布 $N(0,1)$ に法則収束する．ただし $\tilde\theta^*$ は $|\theta_0 - \tilde\theta^*| \leqq |\theta_0 - \tilde\theta|$ を満足する．しかし **4.3.8** から

(12.5.13)
$$g_n'(x_1, \cdots, x_n\,;\,\tilde\theta^*), \qquad g_n'(x_1, \cdots, x_n\,;\,\tilde\theta)$$

は $n \longrightarrow \infty$ のとき $B^*(\theta_0, \theta_0)$ に確率収束する．それゆえ，(12.5.5) 式と同様に次式を得る．

(12.5.14)
$$\lim_{n\to\infty} P\left[\tilde\theta - \frac{\lambda_\gamma}{\sqrt{n}\, g_n'(x_1, \cdots, x_n\,;\,\tilde\theta)} < \theta_0 < \tilde\theta + \frac{\lambda_\gamma}{\sqrt{n}\, g_n'(x_1, \cdots, x_n\,;\,\tilde\theta)}\right] = \gamma.$$

ここに，$\tilde\theta \pm \dfrac{\lambda_\gamma}{\sqrt{n}\, g_n'(x_1, \cdots, x_n\,;\,\tilde\theta)}$ は大きな n に対する θ_0 の漸近的 $100\gamma\%$ 信頼区間の端点で，その長さは $\dfrac{2\lambda_\gamma}{\sqrt{n}\, g_n'(x_1, \cdots, x_n\,;\,\tilde\theta)}$ である．

(c)　漸近的最短信頼区間

(12.5.5) における信頼区間の長さの平方と (12.5.14) のそれとの比 r_n は

(12.5.15)
$$r_n = c_n^2 \cdot d_n^2$$

になる．ただし

(12.5.16)
$$c_n = \frac{g_n'(x_1, \cdots, x_n\,;\,\theta_0)}{h_n'(x_1, \cdots, x_n\,;\,\hat\theta)}$$

(12.5.17)
$$d_n = \frac{g_n'(x_1, \cdots, x_n\,;\,\tilde\theta)}{g_n'(x_1, \cdots, x_n\,;\,\theta_0)}$$

である．さて，c_n, d_n は確率変数である．しかし **4.3.7** から，d_n^2 の分子も分母もともに $B^{*2}(\theta_0, \theta_0)$ に確率収束するから，d_n^2 は 1 に確率収束する．c_n^2 も同様．したがって r_n は

(12.5.18)
$$\frac{B^{*2}(\theta_0, \theta_0)}{B^2(\theta_0, \theta_0)}$$

に確率収束する．これが 1 を越えないことを示そう．

(12.5.7) はすべての n に対して成り立ち，(12.5.8) が使用できる．よって $j=1$ に対して積分記号の下で θ に関して (12.5.7) を微分すると，$\theta = \theta_0$ で

(12.5.19)
$$\int_{R_n} g_n'(x_1, \cdots, x_n\,;\,\theta_0)\,dF_{n0} + \int_{R_n} \sqrt{n}\, g_n(x_1, \cdots, x_n\,;\,\theta_0)\left[\frac{1}{\sqrt{n}} S_n(x_1, \cdots, x_n\,;\,\theta_0)\right] dF_{n0} = 0$$

が得られる．よって，これらの積分の平方は等しい．しかし最初の積分の平方は，$B_n^{*2}(\theta_0, \theta_0)$ である．第 2 項にシュワルツの不等式を適用し，$j=2$ に対する (12.5.7) の積分が 1 であることを用いると

(12.5.20)
$$B_n^{*2}(\theta_0, \theta_0) \leqq \int_{R_n}\left[\frac{1}{\sqrt{n}} S_n(x_1, \cdots, x_n\,;\,\theta_0)\right]^2 dF_{n0}.$$

しかし右辺は n に依存しない $B^2(\theta_0, \theta_0)$ になる．よって
$$B_n^{*2}(\theta_0, \theta_0) \leqq B^2(\theta_0, \theta_0).$$

また $n \longrightarrow \infty$ のときの極限は

$$\frac{B^{*2}(\theta_0,\theta_0)}{B^2(\theta_0,\theta_0)} \leqq 1 \tag{12.5.21}$$

である．よって信頼区間の長さの平方の比は 1 以下の数に確率収束する．(12.5.5) で定義される信頼区間は大きな n に対する θ_0 の漸近的最短 $100\gamma\%$ 信頼区間といわれる．

まとめると次のようになる．

12.5.1　$(x_1,\cdots,x_n)$ を c.d.f. $F(x;\theta_0)$ からの無作為標本とする．ただし，$F(x;\theta)$ はその Ω_0 における θ に対して θ の 2 次導関数に関して正則である．このとき $g_n(x_1,\cdots,x_n;\theta)$ が θ_0 に対する任意の正則推定関数であれば，大きな n に対して θ_0 の漸近的 $100\gamma\%$ 信頼極限は (12.5.14) で与えられる．さらに，(12.5.1) で定義された尤度推定関数 $h_n(x_1,\cdots,x_n;\theta)$ よりも漸近的に短い $100\gamma\%$ 信頼区間を与える θ_0 に対する正則推定関数は存在しない．このように，漸近的最短信頼区間は (12.5.5) で与えられる．

漸近的最短信頼区間問題と推定の漸近的効率の問題との同値性は，公式 (12.3.23) で計算された (12.5.10) における推定量 $\tilde{\theta}$ の漸近的効率が (12.5.21) の左辺の比に他ならないことに注意すればわかるだろう．すなわち (12.5.10) で定義された θ_0 に対する推定量 θ の漸近効率の平方根は次のような比の（確率収束）極限に等しい．ただし，この比は尤度推定関数 $h_n(x_1,\cdots,x_n;\theta)$ によって定まる漸近的 $100\gamma\%$ 信頼区間の長さに対する $g_n(x_1,\cdots,x_n;\theta)$ による対応する信頼区間の長さの比である．

$g_n(x_1,\cdots,x_n;\theta)$ の形が $\sum_{\xi=1}^{n} g(x_\xi;\theta)$ である場合に対する漸近的最短信頼区間の問題は，Wilks (1938 b) によって考察されている．Wald (1942) はさらに一般化している．

12.6　多次元点推定

(a)　はじめの注意

12.2 節から 12.5 節までに得られた結論は，θ が 1 次元の c.d.f. $F(x;\theta)$ からの標本の場合であった．12.1 節のはじめに指摘したように，これらの結果は，標本が，θ が 1 次元の k 次元 c.d.f. $F(x_1,x_2,\cdots,x_k;\theta)$ からの場合でも，記号を若干変えることでそのまま成立する．

さて θ が r 次元で成分 $\theta_1,\cdots,\theta_r$ を持てば，いままでに得られた基本的な結論は読者にとって直感的には明らかでないかも知れないが，r 次元に拡張される．r 次元の場合のすべてについては詳細な証明は与えないが，いくつか結果を述べよう．1 次元パラメータに対する結果に十分に詳しい読者には証明の詳細を補うにはさほど困難はなかろう．

この節を通じて，θ が r 次元で，便宜上 x が 1 次元であるような，c.d.f. $F(x;\theta)$ からの単純無作為標本を考えよう．この節の結果は記号を少し変えるだけで x が k 次元の場合に拡張できる．θ の成分は $(\theta_1,\cdots,\theta_r)$ で，θ の真値 $(\theta_{10},\cdots,\theta_{r0})$ を θ_0 で表わす．1 次元の区間 Ω_0 に対応して，θ_0 を含む r 次元（開）矩形を Ω_{r0} としよう．

θ_0 の推定量 $(\tilde{\theta}_1,\cdots,\tilde{\theta}_r)$ を θ で表わす．θ_0 の<u>不偏推定量</u> $\tilde{\theta}$ は各 $p=1,\cdots,r$ に対して $\tilde{\theta}_p$ が θ_{p0} の不偏推定量となるような推定量である．θ を推定する十分統計量の組は，12.2 (e) 節で定義されている．θ_0 の<u>一致</u>推定量 $\tilde{\theta}$ は $n\longrightarrow\infty$ のとき，$\tilde{\theta}_p$ が θ_{p0} に確率収束するような推定量である．

(b)　多次元推定量の効率

さて効率の概念をどのようにして多次元推定量の場合に拡張するかを見よう．この拡張では，12.1 (b) 節で用いた記号を採用しよう．さらに

$$S_{pn}(x_1,\cdots,x_n;\theta)=\frac{\partial}{\partial\theta_p}\log dF_n(x_1,\cdots,x_n;\theta)=\sum_{\xi=1}^{n}S_p(x_\xi;\theta) \tag{12.6.1}$$

$$S_{pqn}(x_1,\cdots,x_n;\theta)=\frac{\partial}{\partial\theta_q}S_{pn}(x_1,\cdots,x_n;\theta)$$

とする．$\tilde{\theta}$ が Ω_{r0} のある θ の不偏推定量で，c'_p，$p=1,\cdots,r$ が任意の定数ならば

$$\int_{R_n}\sum_{p=1}^{r}(\tilde{\theta}_p-\theta_p)c'_p\,dF_n=0 \tag{12.6.2}$$

と書ける．これはもちろん $\sum_{p=1}^{r}c'_p\tilde{\theta}_p$ が $\sum_{p=1}^{r}c'_p\theta_p$ の不偏推定量であることを示している．成分 $\tilde{\theta}_p$ は 1 次独立であるとしよう．$F(x;\theta)$ はその θ の 1 次偏導関数に関して正則ならば，(12.6.2) を θ_q で微分すると，

$$\int_{R_n}\sum_{p=1}^{r}(\tilde{\theta}_p-\theta_p)c'_p\cdot S_{qn}(x_1,\cdots,x_n;\theta)\,dF_n=c'_q \tag{12.6.3}$$

を得る．ただし，$q=1,\cdots,r$．(12.6.3) の両辺に c_q を掛けて，両辺を q について和をとる．その結果の等式の両辺を 2 乗して，左辺の 2 乗した積分にシュワルツの不等式を

適用すると，Ω_{r0} の θ に対して

(12.6.4)

$$\left(\sum_1^r c_q c_q'\right)^2 \leqslant \left\{\int_{R_n}\left[\sum_1^r (\tilde{\theta}_p-\theta_p)c_p'\right]^2 dF_n\right\}\cdot\left\{\int_{R_n}\left[\sum_1^r S_{qn}(x_1,\cdots,x_n\,;\,\theta)c_q\right]^2 dF_n\right\}$$

を得る．しかし左辺の最初の｛　｝内は $\sigma^2\left(\sum_1^r c_p'\tilde{\theta}_p|\theta\right)$ で第 2 番目の｛　｝は

$n\sigma^2\left[\sum_1^r S_p(x\,;\,\theta)c_p|\theta\right]$ となる．よって $\theta=\theta_0$ では，(12.6.4) は

$$\sigma^2\left(\sum_1^r c_p'\tilde{\theta}_p|\theta_0\right) \geqslant \underset{(c_1,\cdots,c_r)}{\text{l.u.b.}} \frac{\left(\sum_1^r c_p c_p'\right)^2}{n\sigma^2\left[\sum_{p=1}^r S_p(x\,;\,\theta)c_p|\theta_0\right]} \tag{12.6.5}$$

となる．これは $\sum_1^r c_p'\theta_{p0}$ の不偏推定量 $\sum_1^r c_p'\tilde{\theta}_p$ の分散に対する下界を与えている．

(12.6.4) でシュワルツの不等式を適用したところで，等号が成り立つための必要十分条件は

$$K(\theta)\sum_p(\tilde{\theta}_p-\theta_p)c_p' \equiv \sum_p S_{pn}(x_1,\cdots,x_n\,;\,\theta)c_p \tag{12.6.6}$$

なる型の 1 次従属性が x 空間 R_n の確率 0 の集合をのぞいたすべての点で存在することである．ただし $K(\theta)$ は $(x_1,\cdots,x_n)$ には依存せず，組 $c_1,\cdots,c_r$，組 $c_1',\cdots,c_r'$ の両方ともすべては 0 でないとする．(12.6.6) が Ω_{r0} における θ に対して満足されるような θ_0 の不偏推定量が存在すれば，これを**有効推定量**と呼び，$\hat{\theta}$ で表わす．

$\theta=\theta_0$ における $\sqrt{n}\tilde{\theta}_p$，$p=1,\cdots,r$ の共分散行列を $\|B_{pq}'^{(n)}\|$ で表わし，$\theta=\theta_0$ における $S_q/\sqrt{n}$ の共分散行列は 12.1 (b) で定義されているように $\|B_{pq}\|$ で表わす．$\sqrt{n}\tilde{\theta}_p$ と $S_q/\sqrt{n}$ の共分散は (12.6.3) からわかるように δ_{pq} である．ただし δ_{pq} は $p=q$ のとき 1 で，$p\neq q$ のとき 0 である．よって (12.6.4) は

$$\left(\sum_{p,q}B_{pq}'^{(n)}c_p'c_q'\right)\left(\sum_{p,q}B_{pq}c_pc_q\right) \geqslant \left(\sum_{p,q}\delta_{pq}c_pc_q'\right)^2 \tag{12.6.7}$$

と書ける．等号は (12.6.6) が成り立つときのみ成り立つ．もちろん組 $c_1,\cdots,c_r$，組 $c_1',\cdots,c_r'$ は両方ともすべてが 0 でないことを仮定している．

さて，(12.6.7) の不等号が成り立つのは確率変数 $\sqrt{n}\tilde{\theta}_1,\cdots,\sqrt{n}\tilde{\theta}_n, S_1/\sqrt{n},\cdots,S_r/\sqrt{n}$ の共分散行列が正定値のときのみである．このとき，

$$|B_{pq}'^{(n)}|\cdot|B_{pq}| > |\delta_{pq}| = 1 \tag{12.6.8}$$

となる．(12.6.7) の等号が成り立つのは，上の確率変数が (12.6.6) で示された 1 次従属性により，半正定値をなすときのみである．このとき，

$$|B_{pq}'^{(n)}|\cdot|B_{pq}| = 1 \tag{12.6.9}$$

となる．すなわち，(12.6.7) の不等号（等号）は

$$|B_{pq}'^{(n)}|\cdot|B_{pq}| \geqslant 1 \tag{12.6.10}$$

の不等号（等号）が成り立つときにのみ成立する．成分 $\tilde{\theta}_1,\cdots,\tilde{\theta}_r$ は 1 次独立であるので，**3.5.2** から $\|B_{pq}'^{(n)}\|$ は正定値であり，$|B_{pq}'^{(n)}|>0$ となる．同様に $|B_{pq}|>0$．

この結果から，(12.2.15) の一般化として，θ_0 の r 次元不偏推定量 $\tilde{\theta}$ の効率を次式の比で定義できる．

$$\text{eff}\,(\tilde{\theta}|\theta_0) = \frac{1}{|B_{pq}'^{(n)}|\cdot|B_{pq}|}. \tag{12.6.11}$$

有効推定量 $\hat{\theta}$ に対して $\text{eff}\,(\hat{\theta}|\theta_0)=1$ である．

上の結果をまとめると次のようになる．

12.6.1　$(x_1,\cdots,x_n)$ を c.d.f. $F(x\,;\,\theta)$ からの標本とする．ただし θ は r 次元で，$F(x\,;\,\theta)$ は Ω_{r0} において θ の 2 次導関数に関して正則であるとする．θ_0 の不偏推定量で，その成分が 1 次従属でなく，その共分散行列 $\|B_{pq}'^{(n)}\|$ が存在し，かつ正定値であるような不偏推定量 $\tilde{\theta}$ が存在すれば，要素のすべてが 0 ではないような任意の 2 つの定数の組 $c_1,\cdots,c_r$ と $c_1',\cdots,c_r'$ に対して不等式 (12.6.5)，(12.6.10) が成り立つ．さらに両方の等号が成り立つための必要十分条件は $\tilde{\theta}$ が有効推定量となることである．すなわち，(12.6.6) が成り立つことである．

12.7　大標本からの多次元点推定

(a)　スコアの漸近分布

r 次元ベクトル θ に対するスコアは，(12.6.1) で定義したように，成分 $S_{pn}(x_1,\cdots,x_n\,;\,\theta)$，$p=1,\cdots,r$ を持つ r 次元ベクトルである．この成分は，**12.3.1** と同じ条件下

では，大きな n に対して漸近的に r 次元正規分布を持つ．厳密に述べると，

12.7.1　$(x_1, \cdots, x_n)$ を c.d.f. $F(x\,;\theta_0)$ からの標本とする．ただし θ_0 は r 次元．$F(x\,;\theta)$ は Ω_{r0} において θ の 1 次導関数に関して正則であるとする．このとき $\|B_{pq}(\theta,\theta)\|$, $p, q = 1, \cdots, r$ が Ω_{r0} における θ に対して正定値であれば，$(S_{pn}(x_1, \cdots, x_n\,;\theta_0),\ p = 1, \cdots, r)$ は大きな n に対して漸近的に r 次元分布 $N(\{0\}, \|nB_{pq}\|)$ に従って分布する．ただし $B_{pq} = B_{pq}(\theta_0, \theta_0)$.

この証明は **12.3.1** で示した 1 次元の場合と同様なので，読者に演習問題として残しておく．

(b)　最尤推定量の収束

1 次元の場合のように，各 n に対して成分が (12.6.6) を満足する θ_0 の不偏推定量 $\tilde{\theta}$ が存在しないならば，ある条件下では，確率 1 で θ_0 に収束する方程式

(12.7.1)　$$S_p(x_1, \cdots, x_n\,;\theta) = 0, \qquad p = 1, \cdots, r$$

の解の列が存在する．

厳密に述べれば，次のような **12.3.2** の r 次元への拡張が得られる．

12.7.2　$(x_1, \cdots, x_n)$ を c.d.f. $F(x\,;\theta_0)$ からの標本とする．ただし，$\theta_0 = (\theta_{10}, \cdots, \theta_{r0})$ は r 次元で，$F(x\,;\theta)$ は Ω_{r0} において θ の 1 次導関数に関して正則である．$S_p(x\,;\theta)$, $p = 1, \cdots, r$ を R_1 の確率 0 の集合をのぞくすべての x の値に対して Ω_{r0} で θ の連続関数とする．このとき，$(\theta_{10}, \cdots, \theta_{r0})$ に概収束する (12.7.1) の解の列が存在する．いま解がある n_0 より大きな n に対して一意のベクトル $(\hat{\theta}_1, \cdots, \hat{\theta}_r)$ であれば，$n \longrightarrow \infty$ のとき，このベクトル列は $(\theta_{10}, \cdots, \theta_{r0})$ に概収束する．

12.7.2 の証明は **12.3.2** の証明をそのまま拡張すればよいので省略する．

(c)　最尤推定量の漸近分布

12.3.3 の r 次元への拡張は次のように述べられる．

12.7.3　$(x_1, \cdots, x_n)$ を c.d.f. $F(x\,;\theta_0)$ からの標本とする．ここで θ_0 は r 次元，$F(x\,;\theta)$ は Ω_{r0} において θ の 1 次および 2 次導関数に関して正則であるとする．このとき，(12.7.1) を満たす最尤推定量 $(\hat{\theta}_1, \cdots, \hat{\theta}_r)$ が，ある n_0 より大きな n に対して一意で，$\prod_{\xi=1}^{n} F(x_\xi\,;\theta)$ に関して可測ならば，最尤推定量は大きな n に対して r 次元正規分布 $N(\{\theta_{p0}\}, \|nB_{pq}\|^{-1})$ に漸近分布する．ただし $\|B_{pq}\| = \|B_{pq}(\theta_0, \theta_0)\|$.

12.7.3 の論理の流れは **12.3.3** の場合と同じであるので省略する．

(d)　最尤推定量の漸近効率

r 次元推定量において，$\tilde{\theta}$ の各成分が θ_0 の対応する各成分の一致推定量ならば，$\tilde{\theta}$ は θ_0 の一致推定量であるという．

$\tilde{\theta}$ が θ_0 の一致推定量で，成分 $(\tilde{\theta}_1, \cdots, \tilde{\theta}_r)$ については，

$$(\sqrt{n}(\tilde{\theta}_1 - \theta_0), \cdots, \sqrt{n}(\tilde{\theta}_r - \theta_0))$$

が $n \longrightarrow \infty$ のとき，極限分布として r 次元正規分布 $N(\{0\}, \|B^*_{pq}\|^{-1})$ を持つとする．このとき

(12.7.2)　$$\mathrm{leff}\,(\tilde{\theta}|\theta_0) = \frac{|B^*_{pq}|}{|B_{pq}|}$$

によって $\tilde{\theta}$ の漸近効率を定義する．

(12.7.3)　$$\lim_{n\to\infty} \sigma^2(\sqrt{n}\sum_p c'_p\theta_p) = \sum_{p,q} B^{pq}c'_pc'_q$$

である．

12.7.3 から

$$(\sqrt{n}(\hat{\theta}_1 - \theta_{10}), \cdots, \sqrt{n}(\hat{\theta}_r - \theta_{r0}))$$

は $n \longrightarrow \infty$ のとき極限分布として $N(\{0\}, \|B_{pq}\|^{-1})$ を持つ．よって

(12.7.4)　$$\mathrm{leff}\,(\hat{\theta}|\theta_0) = \frac{|B_{pq}|}{|B_{pq}|} = 1$$

である．ゆえに

12.7.4　**12.7.3** の仮定の下では，最尤推定量 $\hat{\theta}$ は θ_0 の推定に対する漸近効率 1 を持つ．

1 次元パラメータの有効推定量の場合と同様に，多次元パラメータの有効推定量の最初の研究は Fisher (1922) によってなされた．近代的な数学的モデルに立脚した最近の研

究は Barankin と Gurland (1951) がある．

12.8 多次元信頼領域

12.4 節で論じたように，1 次元パラメータ θ_0 の区間推定においては，本質的な考え方は固定された信頼係数 γ および c.d.f. $F(x;\theta)$ からの標本 $(x_1,\cdots,x_n)$ に対して

$$P(\underline{\theta}<\theta<\bar{\theta}|\theta)=\gamma$$

なる確率変数 $\underline{\theta}<\bar{\theta}$ が存在することである．ある条件の下で $\underline{\theta}$, $\bar{\theta}$ を構成する方法は **12.4.1** に示した．12.4 (b) で簡単に指摘したように，**12.4.1** において Ω 上で θ について $g_n(x_1,\cdots,x_n;\theta)$ の単調性をはずせば，次を満たす R_1 の確率集合 $E_1(x_1,\cdots,x_n)$ が存在する．これを単に E_1 で表わす．すなわち E_1 は

(12.8.1) $$g_1<g_n(x_1,\cdots,x_n;y)<g_2$$

なる Ω のすべての実数 y からなっていて，

(12.8.2) $$P(g_1<g_n(x_1,\cdots,x_n;\theta_0)<g_2|\theta_0)=P(\theta_0\in E_1|\theta_0)=\gamma$$

なる θ_0 に依存しない．いい換えれば，E_1 が θ_0 を含む確率は γ である．Ω における θ のすべての値に対して $g(x_1,\cdots,x_n;\theta_0)$ が確率変数だから，E_1 はボレル集合である．すなわち E_1 が θ を含む確率が θ に依存しないのは，$g_n(x_1,\cdots,x_n;\theta)$ の c.d.f. が θ に依存しないからである．特に，E_1 は θ_0 を含む確率が γ になる Ω の確率集合である．ただし，ここでは標本$(x_1,\cdots,x_n)$ は $F(x;\theta_0)$ から抽出されているとしている．

ゆえに (12.8.2) は区間推定の集合推定への拡張を与えている．この拡張は次に示すように，信頼区間推定を r 次元パラメータの場合にまで一般化できることを示唆している．

12.8.1 $(x_1,\cdots,x_n)$ を c.d.f. $F(x;\theta_0)$ からの標本とする．ここで θ_0 は r 次元．$g_n(x_1,\cdots,x_n;\theta)$ は確率 0 の集合をのぞいた標本空間 R_n のすべての点とパラメータ空間 Ω_r のすべての点 θ で定義されていて，c.d.f. が $F(x;\theta)$ であるとき，g の c.d.f. が θ に依存しないような確率変数とする．また $E_r(x_1,\cdots,x_n)=E_r$ を $g_1<g_n(x_1,\cdots,x_n;y)<g_2$ なる Ω_r の点 $y=(y_1,\cdots,y_r)$ の集合とする．ここで g_1, g_2 は

(12.8.3) $$P(g_1<g_n(x_1,\cdots,x_n;\theta)<g_2|\theta)=\gamma$$

なるように選ばれている．ただし γ は θ に依存しないとする．このとき E_r は真のパラメータ点 θ_0 に対して

(12.8.4) $$P(\theta_0\in E_r|\theta_0)=\gamma$$

となるような r 次元確率集合である．

この定理の証明は容易なので省略する．集合 E_r を θ_0 に対する $100\gamma\%$ 信頼領域と呼ぶ．k 次元パラメータを推定する信頼領域の例は 10.4 節に示してある．

一般に $g_n(x_1,\cdots,x_n;\theta)$ が，母集団 c.d.f. が $F(x;\theta)$ のとき，θ に依存しない c.d.f. を持つ成分 $g_{p'n}(x_1,\cdots,x_n;\theta)$, $p'=1,\cdots,r'$ からなるベクトル関数であれば，

(12.8.5) $$P((g_{1n},\cdots,g_{r'n})\in E^*_{r'})=\gamma$$

なる $R_{r'}$ における任意の集合 $E^*_{r'}$ に対して

(12.8.6) $$(g_{1n}(x_1,\cdots,x_n;y),\cdots,g_{r'n}(x_1,\cdots,x_n;y))\in E^*_{r'}$$

なる Ω_r のすべての点 $y=(y_1,\cdots,y_r)$ からなる集合 E_r は，θ_0 に対する $100\gamma\%$ 信頼域である．すなわち

(12.8.7) $$P(\theta_0\in E_r)=\gamma.$$

通常，最もよく役に立ち，しかも重要な関数 $g_n(x_1,\cdots,x_n;\theta)$ は凸な，あるいは少なくとも単連結な信頼領域（確率集合）E_r を与えるものである．このような集合は表現が容易であるうえに，用い易いからである．$r=r'=1$ のときは，$g_n(x_1,\cdots,x_n;\theta)$ が **12.4.1** で示したように θ の連続かつ単調な関数ならば，E_1 は上述の性質を持つ．r, r' が一般的な値のときは，この結果はさらに複雑になる．しかし，E_r に必要とされるものは凸性あるいは単連結性の他には，ある意味における“最小性”であることは直感的に明らかである．しかしながら，これは小さな n に対しては複雑な問題であるが，n が大きいときは，ある条件の下では満足すべき漸近的解がある．それを次節で述べよう．

例題 前述の考え方の例を示そう．$(x_1,\cdots,x_n)$ を $N(\mu_0,\sigma_0^2)$ からの標本とする．このとき，**8.4.1** より，関数

$$g(x_1,\cdots,x_n;\mu_0,\sigma_0)=\frac{(n-1)s^2+n(\bar{x}-\mu_0)^2}{\sigma_0^2}$$

はカイ 2 乗分布 $C(n)$ を持つ．よって与えられた γ に対して

$$P(g<\chi^2_\gamma)=\gamma$$

なる χ^2_γ を求めることができる．y_1y_2 平面の $y_2>0$ なる半平面における集合で，式

$$\frac{y_2^2}{n}-\frac{(y_1-\bar{x})^2}{\chi^2_\gamma}=\frac{(n-1)s^2}{n\chi^2_\gamma}$$

で表わされる 2 つの双曲線間の領域を E_2 で表わす．(μ_0, σ_0) が真のパラメータの点ならば

$$P((\mu_0, \sigma_0) \in E_2) = \gamma$$

がわかる．

さて 2 次元のベクトル関数

$$g_{p'}(x_1, \cdots, x_n ; \mu_0, \sigma_0), \qquad p' = 1, 2$$

を考えよう．ただし

$$g_1 = \frac{n(\bar{x} - \mu_0)^2}{\sigma_0^2}, \qquad g_2 = \frac{n-1}{\sigma_0^2} s^2.$$

8.4.2 より，g_1, g_2 はカイ 2 乗分布 $C(1)$, $C(n-1)$ に従って独立に分布する．

$$P(g_1 < \chi_1^2, g_2 > \chi_2^2) = \gamma$$

となるように χ_1^2, χ_2^2 を選び，E_2^* を

$$\frac{n(y_1 - \bar{x})^2}{y_2} < \chi_1^2, \qquad \frac{\sqrt{(n-1)}s}{y_2} > \chi_2^2$$

なる $y_1 y_2$ 平面の領域とすれば

$$P((\mu_0, \sigma_0) \in E_2^*) = \gamma$$

を得る．ただし $y_2 > 0$.

このように，この例では E_2, E_2^* はそれぞれ 1 つ，2 つの成分を持つベクトル推定関数により決まる信頼領域である．直感的に E_2^* は E_2 よりも満足すべきものである．なぜなら，E_2^* は有界で凸であるが，E_2 はこれらの性質を持っていないからである．もし $g(x_1, \cdots, x_n ; \mu_0, \sigma_0)$ として $(\bar{x}, s^2)$ の p.d.f. を選べば，"最小性" という基準のもとでは，E_2^* よりも満足できる E_2^{**} を求めることができる．しかし E_2^{**} は E_2^* よりも複雑で，その詳細を記述することはできないだろう．n が大きな場合には，(μ_0, σ_0) を推定する漸近的最小領域を求める最適な方法が存在する．漸近的最小信頼領域の一般的な問題は次節で取り上げる．

12.9　大標本からの漸近的最小信頼領域

ここでは **12.5.1** の r 次元への拡張を考えよう．

12.7.1 と **9.3.2** より，θ の真値が θ_0 ならば，確率変数列

(12.9.1)
$$U_n(x_1, \cdots, x_n ; \theta_0) = n \sum_{p,q=1}^{r} B^{pq} h_{pn}(\theta_0) h_{qn}(\theta_0), \qquad n = 1, 2, \cdots$$

はカイ 2 乗分布 $C(r)$ に法則収束する．ただし (12.6.1) で定義したように，$h_{pn}(\theta) = \dfrac{1}{n} S_{pn}(x_1, \cdots, x_n ; \theta)$ である．**12.7.3** の強い条件の下では，U_n に微分の平均値の定理を

適用でき，

$$U_n^* \equiv U_n, \qquad \hat{\theta} \text{ の値が } \Omega_{r0} \text{ に属すとき}$$

$$U_n^* \equiv 0, \qquad \hat{\theta} \text{ の値が } \Omega_{r0} \text{ に属さないとき}$$

と書ける．ただし，確率 0 の集合をのぞいた標本空間 R_n のすべての点で，

(12.9.2)
$$U_n^*(x_1, \cdots, x_n ; \theta_0) = \sum_{p,q} B_{pq}^{(n)} [\sqrt{n}(\theta_{p0} - \hat{\theta}_p)][\sqrt{n}(\theta_{q0} - \hat{\theta}_q)]$$

である．ここに

(12.9.3)
$$B_{pq}^{(n)} = \Big[\sum_{p',q'=1}^{r} B^{p'q'} \Big(\frac{\partial h_{p'n}}{\partial \theta_p}\Big)\Big(\frac{\partial h_{q'n}}{\partial \theta_q}\Big)\Big]_{\theta = \hat{\theta}^*}$$

$$|\theta_{p0} - \hat{\theta}_p^*| < |\theta_{p0} - \hat{\theta}_p|, \qquad p = 1, \cdots, r.$$

12.7.3 の条件と，**4.3.7**, **4.3.8** から，$B_{pq}^{(n)}$ は $n \longrightarrow \infty$ のとき $\sum_{p',q'=1}^{r} B^{p'q'} B_{p'p} B_{q'q}$ $(= B_{pq})$ に確率収束する．したがって，U_n, U_n^*, $n = 1, 2, \cdots$ はともにカイ 2 乗分布 $C(r)$ に法則収束する確率変数列であることが確かめられる．さて，χ_γ^2 を次のように選ぼう．

(12.9.4)
$$\lim_{n \to \infty} P(U_n < \chi_\gamma^2) = \lim_{n \to \infty} P(U_n^* < \chi_\gamma^2) = \gamma.$$

また $E_r(x_1, \cdots, x_n)$, $E_r^*(x_1, \cdots, x_n)$ を Ω_{r0} における，それぞれ

(12.9.5)
$$U_n(x_1, \cdots, x_n ; y) < \chi_\gamma^2$$

(12.9.6)
$$U_n^*(x_1, \cdots, x_n ; y) < \chi_\gamma^2$$

を満足する $y = (y_1, \cdots, y_r)$ の点集合とする．それゆえ，θ_0 が θ の真値ならば，

(12.9.7)
$$\lim_{n \to \infty} P(\theta_0 \in E_r) = \lim_{n \to \infty} P(\theta_0 \in E_r^*) = \gamma.$$

すなわち E_r と E_r^* は θ_0 の推定に対して漸近的に同値な信頼領域である．しかし E_r^* は θ_0 を中心に持つ Ω_{r0} における楕円体で，この体積は大きな n に対して近似的に

(12.9.8)
$$\frac{K(r)\chi_\gamma}{\sqrt{n}\,|B_{pq}^{(n)}|}$$

である．ここで $K(r)$ は r にのみ依存する定数である．

さて $g_n(x_1, \cdots, x_n ; \theta)$ を 1 次独立な成分 $g_{pn}(x_1, \cdots, x_n ; \theta)$, $p = 1, \cdots, r$ を持ち，成分の θ の 1 次導関数が

(12.9.9)
$$g_{pqn}(x_1, \cdots, x_n ; \theta) = (\partial/\partial\theta_q) g_{pn}(x_1, \cdots, x_n ; \theta), \qquad p, q = 1, \cdots, r$$

であるようなベクトル確率変数とする．$g_{pn}(x_1, \cdots, x_n ; \theta)$, $g_{pqn}(x_1, \cdots, x_n ; \theta)$ をそれ

それぞれ $g_p(\theta)$, $g_{pq}(\theta)$ で表わす．$g_p(\theta)$, $p=1,\cdots,r$ と $g_{pq}(\theta)$, $p,q=1,\cdots,r$ は (12.5.7) でそれぞれ $g_n(x_1,\cdots,x_n;\theta)$, $g'_n(x_1,\cdots,x_n;\theta)$ に対して示した条件 (iii) から (v) を満足するものと仮定する．また対応する関数 A_n^*, B_n^* を $A_{pn}^*(\theta_0,\theta)$, $p=1,\cdots,r$ と $B_{pqn}^*(\theta_0,\theta)$, $p,q=1,\cdots,r$ で表わそう．ただし (12.5.8) は単に置き換えるだけで成立するとする．$B_{pq}^*(\theta,\theta)=\lim_{n\to\infty}B_{pqn}^*(\theta,\theta)$, $B_{pq}^*=B_{pq}^*(\theta_0,\theta_0)$ とする．このとき，(12.5.7) の (i) に対応して，$j=1$ のとき

(12.9.10)
$$\int_{R_n}\sqrt{n}\,g_p(\theta)\,dF_n=0.$$

$j=2$ のとき

(12.9.11)
$$\int_{R_n}[\sqrt{n}\,g_p(\theta)][\sqrt{n}\,g_q(\theta)]\,dF_n=C_{pqn}(\theta)$$

となる．また (12.5.7) の (ii) に対応して，母集団分布において $\theta=\theta_0$ ならば，$(\sqrt{n}\,g_p(\theta_0),\ p=1,\cdots,r)$ は大きな n に対して分布 $N(\{0\},\|C_{pq}\|)$ に従って漸近的に分布することを仮定する．ただし

(12.9.12)
$$C_{pq}=\lim_{n\to\infty}C_{pqn}(\theta_0)$$

で $\|C_{pq}\|$ は正定値であるとする．

上述の条件を満足するベクトル関数 $g(x_1,\cdots,x_n;\theta)$ は θ_0 に対する**正則推定ベクトル関数**と呼ばれる．読者は尤度推定ベクトル関数 $h_p(\theta)=\dfrac{1}{n}S_{pn}(x_1,\cdots,x_n;\theta)$ が正則推定ベクトル関数になることに気づくであろう．

このような推定関数に対して，$n\geqq$(ある n_0) のとき，方程式

(12.9.13)
$$g_{pn}(x_1,\cdots,x_n;\theta)=0,\qquad p=1,\cdots,r,\quad n=n_0,n_0+1,\cdots$$

の一意解の列 $(\tilde\theta_1,\cdots,\tilde\theta_r)$ が存在して，θ_0 が θ の真値のとき，確率変数列 $[\sqrt{n}(\tilde\theta_1-\theta_{10}),\cdots,\sqrt{n}(\tilde\theta_r-\theta_{r0})]$, $n=n_0,n_0+1,\cdots$ が $n\longrightarrow\infty$ のとき r 次元分布 $N(\{0\},\|D^{pq}\|)$ を持つことが示される．ただし

(12.9.14)
$$D_{pq}=\sum_{p'q'=1}^{r}C^{p'q'}B_{p'p}^*B_{q'q}^*.$$

9.3.2 より，θ_0 が θ の真値ならば，

(12.9.15)
$$V_n=n\sum_{p,q=1}^{r}C^{pq}g_p(\theta_0)g_q(\theta_0)$$

で定義される $V_n(x_1,\cdots,x_n;\theta_0)$ は $n\longrightarrow\infty$ のとき極限分布としてカイ2乗分布 $C(r)$ を持つ．

さらに，θ_0 が θ の真値ならば，$n\longrightarrow\infty$ のとき確率変数 U_n, U_n^* のおのおのがともにカイ2乗分布 $C(r)$ を極限分布として持つことを示したのと同じ方法で，V_n と次式で定義される確率変数 $V_n^*(x_1,\cdots,x_n;\theta_0)$ が $n\longrightarrow\infty$ のときともにカイ2乗分布 $C(r)$ に法則収束することが示される．

(12.9.16)
$$V_n^*=\sum_{p,q=1}^{r}D_{pq}^{(n)}[\sqrt{n}(\tilde\theta_p-\theta_{p0})][\sqrt{n}(\tilde\theta_q-\theta_{q0})].$$

ただし

(12.9.17)
$$D_{pq}^{(n)}=\sum_{p',q'=1}^{r}C^{p'q'}g_{p'p}(\tilde\theta^*)g_{q'q}(\tilde\theta^*).$$

ここに

$$|\theta_{p0}-\tilde\theta_p^*|<|\theta_{p0}-\tilde\theta_p|,\qquad p=1,\cdots,r.$$

集合 $E'_r(x_1,\cdots,x_n)$, $E_r^{*\prime}(x_1,\cdots,x_n)$ がそれぞれ

(12.9.18)
$$V_n(x_1,\cdots,x_n;y)<\chi_\Upsilon^2$$

(12.9.19)
$$V_n^*(x_1,\cdots,x_n;y)<\chi_\Upsilon^2$$

なる Ω_{r0} の点 $y=(y_1,\cdots,y_r)$ の集合として定義されているとき，両者は Ω_{r0} における漸近的に同値な集合である．すなわち $n\longrightarrow\infty$ のとき，$P(E'_r)$, $P(E_r^{*\prime})$ の極限は等しい．ここで χ_Υ^2 は (12.9.4) のように選ばれている．しかし $E_r^{*\prime}$ は体積

(12.9.20)
$$\frac{K(r)\chi_\Upsilon}{\sqrt{n|D_{pq}^{(n)}|}}$$

を持つ中心 θ_0 の楕円体である．さて，(12.9.8) で与えられる E_r^* の体積と，(12.9.20) で与えられる $E_r^{*\prime}$ の体積との比 r_n は $n\longrightarrow\infty$ のとき

(12.9.21)
$$r=\frac{\sqrt{|D_{pq}|}}{\sqrt{|B_{pq}|}}$$

で与えられる r に（確率）収束する．しかし，(12.9.14) で与えられた D_{pq} を代入すると

(12.9.22)
$$r=\frac{|B_{pq}^*|}{\sqrt{|C_{pq}|\cdot|B_{pq}|}}$$

となる．

$g_p(\theta)$ に対する仮定より，(12.9.10) を積分記号下で θ_q に関して微分できて，

(12.9.23)
$$\int_{R_n} g_{pq}(\theta)\,dF_n + \int_{R_n} g_p(\theta)S_q(\theta)\,dF_n = 0$$

を得る．ここで $S_q(\theta)$ は $S_{qn}(x_1,\cdots,x_n\,;\,\theta)$ を表わす．

(12.9.23) より，$\theta=\theta_0$ において

(12.9.24)
$$\int_{R_n} g_{pq}(\theta_0)\,dF_n = -\operatorname{cov}\left(\sqrt{n}\,g_p(\theta_0), \frac{S_q(\theta_0)}{\sqrt{n}}\right)$$

となる．しかし (12.9.24) の左辺は $B^*_{pqn}(\theta_0,\theta_0)$．ゆえに右辺もまた $B^*_{pqn}(\theta_0,\theta_0)$．これを略して B^*_{pqn} で表わす．この結果，$2r$ 次元確率変数列

$$\left(\sqrt{n}\,g_1(\theta_0),\cdots,\sqrt{n}\,g_r(\theta_0),\frac{1}{\sqrt{n}}S_1(\theta_0),\cdots,\frac{1}{\sqrt{n}}S_r(\theta_0)\right),\qquad n=1,2,\cdots$$

は次の共分散行列を持つ $2r$ 次元確率変数に確率収束する．

(12.9.25)
$$\left\|\begin{array}{c:c} C_{pq} & B^*_{pq} \\ \hdashline B^*_{pq} & B_{pq}\end{array}\right\|.$$

さらに，(12.9.25) は共分散行列だから

(12.9.26)
$$|C_{pq}|\cdot|B_{pq}| \geqq |B^*_{pq}|^2$$

となる．よって (12.9.22) で与えられた信頼領域の体積比の（確率収束）極限は 1 を越えない．

E_r, E'_r はそれぞれ E^*_r と $E^{*\prime}_r$ に漸近的に同値である．これは次の事柄を意味している．すなわち，(12.9.18) の $E'_r(x_1,\cdots,x_n)$ によって定義された θ_0 の $100\gamma\%$ 信頼領域が，$g_p(\theta)$ を $h_p(\theta)$ で置き換えて同様にして得られる θ_0 の $100\gamma\%$ 信頼領域よりも，漸近的に小さくなるような正則推定ベクトル関数 $g(x_1,\cdots,x_n\,;\,\theta)$ は存在しない．

上の議論の結果を要約すると，次のような **12.5.1** の r 次元への拡張が得られる．

12.9.1　$(x_1,\cdots,x_n)$ を c.d.f. $F(x\,;\,\theta_0)$ からの標本とする．ただし θ_0 は r 次元で，$F(x\,;\,\theta)$ は Ω_{r0} において，θ の 1 次，および 2 次導関数のすべてに関して正則である．このとき，$g_{pn}(x_1,\cdots,x_n\,;\,\theta)$, $p=1,\cdots,r$ が θ_0 の任意の正則推定ベクトル関数ならば，θ_0 の漸近的 $100\gamma\%$ 信頼領域は (12.9.18) を満たす点 $y=(y_1,\cdots,y_r)$ で定義された集合 E'_r によって与えられる．さらに (12.9.1) で定義された尤度推定ベクトル関数 $h_{pn}(x_1,\cdots,x_n\,;\,\theta)$, $p=1,\cdots,r$ を用いて得られた θ_0 の漸近的 $100\gamma\%$ 信頼領域よりも小さい漸近的 $100\gamma\%$ 信頼領域を与える正則推定ベクトル関数は存在しない．よって漸近的最小 $100\gamma\%$ 信頼領域 E_r は (12.9.5) を満たす点 $y=(y_1,\cdots,y_r)$ からなっている．

多次元パラメータの場合の漸近的最小信頼領域の問題は Bartlett(1953), Beale(1960), Wilks と Daly (1939) に論じられている．さらに進んだ研究に興味を持つ読者はこれらの文献を参照されたい．

例題　漸近的最小 $100\gamma\%$ 信頼領域を求める例題として k パラメータ p.f.

$$dF(s\,;\,\theta) = \theta_1^{s_1}\cdots\theta_{k+1}^{s_{k+1}}$$

を持つ多項母集団からの標本を考えよう．ここで確率変数 $(s_1,\cdots,s_{k+1})$ は $k+1$ 個の値 $(1,0,\cdots,0),\cdots,(0,\cdots,0,1)$ の 1 つそしてただ 1 つのみをとることができる．ただし $\theta_1>0,\cdots,\theta_{k+1}>0$ で $\theta_1+\cdots+\theta_{k+1}=1$ である．便宜上 $\theta_1,\cdots,\theta_k$ を信頼領域から推定される k 個のパラメータとしよう．さて，$(s_{1\xi},\cdots,s_{k+1\xi},\ \xi=1,\cdots,n)$ をこの分布からの標本とする．$h_p(\theta)$ に対して

$$h_p(\theta) = \frac{1}{n}\sum_{\xi=1}^{n}\frac{\partial\log dF(s_\xi\,;\theta)}{\partial\theta_p} = \frac{1}{n}\sum_{\xi=1}^{n}\left(\frac{s_{p\xi}}{\theta_p}-\frac{s_{k+1\xi}}{\theta_{k+1}}\right)$$

を得る．ここで $\theta_{k+1}=1-\theta_1-\cdots-\theta_k$．$\sum_{\xi=1}^{n}s_{p\xi}=x_p$ と置くと

$$h_p(\theta) = \frac{x_p}{n\theta_p}-\frac{x_{k+1}}{n\theta_{k+1}},\qquad p=1,\cdots,k$$

と書ける．ただし $x_{k+1}=n-x_1-\cdots-x_k$．さらに

$$B_{pq} = \mathscr{E}\left[\left(\frac{\partial\log dF(s\,;\,\theta)}{\partial\theta_p}\right)\left(\frac{\partial\log dF(s\,;\,\theta)}{\partial\theta_q}\right)\right] = \mathscr{E}\left[\left(\frac{s_p}{\theta_p}-\frac{s_{k+1}}{\theta_{k+1}}\right)\left(\frac{s_q}{\theta_q}-\frac{s_{k+1}}{\theta_{k+1}}\right)\right]$$
$$= \frac{\delta_{pq}}{\theta_p}+\frac{1}{\theta_{k+1}},\qquad p,q=1,\cdots,k.$$

ここで δ_{pq} はクロネッカーの δ，すなわち $p=q$ なら 1，$p\neq q$ なら 0 である．逆行列の要素 B^{pq} は $B^{pq}=\delta_{pq}\theta_p-\theta_p\theta_q$ となる．ゆえに (12.9.1) の 2 次形式 U_n は代数的に変形すると

$$U_n = n\sum_{p,q=1}^{k}B^{pq}h_p(\theta)h_q(\theta) = \sum_{p=1}^{k+1}\frac{(x_p-n\theta_p)^2}{n\theta_p}$$

となる．これは K. Pearson (1900) が最初に示した古典的な χ^2 関数である．よって，大きな n に対して k 次元パラメータ $\theta=(\theta_1,\cdots,\theta_k)$ を推定する $100\gamma\%$ 漸近的最小信頼領域は

$$\sum_{p=1}^{k+1}\frac{(x_p-ny_p)^2}{ny_p} < \chi^2_\gamma$$

なる正の実ベクトル $x=(y_1,\cdots,y_k)$ の集合 E_k によって与えられる．ただし $y_{k+1}=1-y_1-\cdots-y_k$ で，χ^2_γ はカイ 2 乗分布 $C(k)$ の $100\gamma\%$ の点である．

問　題

12.1 p.d.f.

$$f_1(x\,;\theta)=\begin{cases}\dfrac{1}{\theta}, & 0<x\leqslant\theta\\ 0, & \text{その他}\end{cases}$$

$$f_2(x\,;\theta')=\begin{cases}\dfrac{1}{\theta'}, & 0<x\leqslant\theta'\\ 0, & \text{その他}\end{cases}$$

を持つ 2 つの分布は $\theta'=\theta$ でなければ互いに絶対連続にならないことを示せ．

12.2 x を c.d.f.

$$F(x\,;\theta)=\begin{cases}0, & x\leqslant 0\\ 1-e^{-\theta x}, & x>0\end{cases}$$

を持つパラメトリック分布に従う確率変数とする．ただし θ は $(0,+\infty)$ 上の任意の数．$\mathscr{E}(S(x\,;\theta))=0$ を示し，$\sigma^2(S(x\,;\theta))$ を求めよ．また $H(\theta,\theta')$ を求め，これは固定した θ に対して $\theta'=\theta$ のとき，θ' に関して最大になることを示せ．

12.3 パラメトリック p.f.

$$p(x\,;\theta)=(1-\theta)\theta^{x-1},\qquad x=1,2,\cdots$$

を持つ離散型確率変数 x がある．ただし θ は $(0,1)$ 上の任意の点．このとき $\mathscr{E}(S(x\,;\theta))=0$ を示せ．また $\sigma^2(S(x\,;\theta))$ を求めよ．

12.4 $(x_1,\cdots,x_n)$ が有限な k 次モーメント $\mathscr{E}(x^k)$ を持つ任意の分布からの標本ならば，$\dfrac{1}{n}\displaystyle\sum_{\xi=1}^{n}x_\xi^k$ は $\mathscr{E}(x^k)$ の一致推定量であることを示せ．

12.5 $\bar{x}$ が 2 項分布 $Bi(1,p)$ からの標本平均ならば，$\bar{x}$ は p の十分推定量であることを示せ．さらに $\bar{x}$ は p の有効推定量であることを示せ．

12.6 $(x_1,\cdots,x_n)$ は p.d.f.

$$\theta e^{-\theta x},\qquad x>0$$

からの標本とする．ただし θ は正のパラメータ．このとき，標本平均 $\bar{x}$ は θ の十分推定量で，$(n-1)/n\bar{x}$ は $\bar{x}$ に依存する θ の唯一の不偏推定量であることを示せ．

12.7 (続き) $\bar{x}$ は $1/\theta$ の最尤推定量で，θ の不偏推定量よりも分散が大きいことを示せ．

12.8 $\bar{x}$, s^2 を正規分布 $N(\mu,\sigma^2)$ からの大きさ $n>1$ の標本の標本平均，標本分散とする．このとき 2 次元確率変数 $(\bar{x},s^2)$ はベクトルパラメータ (μ,σ^2) の十分統計量であることを示せ．

12.9 $\hat{\theta}_1$ と $\hat{\theta}_2$ がパラメータ θ の 2 つの有効推定量のとき，$\hat{\theta}_1$ と $\hat{\theta}_2$ の相関係数は 1 であることを示せ．

12.10 $\hat{\theta}$ を θ の有効推定量，$\tilde{\theta}$ を θ の他の任意の不偏推定量とし，それぞれ $\sigma^2, k\sigma^2$, $k>1$ が $\hat{\theta}$ と $\tilde{\theta}$ の分散であるとき，$\hat{\theta}$ と $\tilde{\theta}$ の相関係数は $1/\sqrt{k}$ に等しいことを示せ．

12.11 (続き) $\hat{\theta}$ が θ の有効推定量であり，$\tilde{\theta}_1,\tilde{\theta}_2$ が等分散 $k\sigma^2$, $k>1$ を持つ，有効ではないが不偏推定量であるとき，$\tilde{\theta}_1$ と $\tilde{\theta}_2$ の相関係数は少なくとも $(2-k)/k$ であることを示せ．

12.12 $(x_1,\cdots,x_k)$ を $\theta\in\Omega_0$ に対して p.d.f. $f(x_1,\cdots,x_k\,;\theta)$ を持つ k 次元確率変数とする．また $g(x_1,\cdots,x_n)$ を $\phi(\theta)$ の不偏推定量とする．このとき f が Ω_0 において θ の 1 次導関数に関して正則で，$\phi(\theta)$ が Ω_0 において θ に関する導関数を持てば

$$\sigma^2(g)\geqslant\left(\frac{\partial\phi}{\partial\theta}\right)^2\Big/\mathscr{E}\left(\frac{\partial\log f}{\partial\theta}\right)^2$$

になることを示せ．また等号が成り立つための必要十分条件は確率 1 で $\partial\log f/\partial\theta\equiv C(g-\phi(\theta))$ のときであることを示せ．ただし C は $(x_1,\cdots,x_k)$ に依存しない．

12.13 $(x_{(1)},\cdots,x_{(n)})$ は矩形分布 $R(\theta/2,\theta)$ からの標本の順序統計量である．このとき，$x_{(n)}$ は θ の十分推定量となり，$(n-1)x_{(n)}/n$ は θ の有効推定量となることを示せ．また $(x_{(n)},x_{(n)}/\sqrt[n]{1-\gamma})$ は θ の $100\gamma\%$ 信頼区間であることを示せ．

12.14 $(x_{(1)},\cdots,x_{(n)})$ は矩形分布 $R((\theta_1+\theta_2)/2,\theta_2-\theta_1)$, $\theta_2>\theta_1>0$ からの標本の順序統計量である．このとき 2 次元確率変数 $(x_{(1)},x_{(n)})$ は (θ_1,θ_2) の十分推定量になることを示せ．

$$\frac{nx_{(1)}}{n-1}-\frac{x_{(n)}}{n-1},\qquad \frac{nx_{(n)}}{n-1}-\frac{x_{(1)}}{n-1}$$

はそれぞれ θ_1, θ_2 の最小分散推定量となることを示せ．また

$$\frac{x_{(1)}+x_{(n)}}{2},\qquad \frac{n+1}{n-1}(x_{(n)}-x_{(1)})$$

は中点 $(\theta_1+\theta_2)/2$ と範囲 $\theta_2-\theta_1$ の最小分散推定量となることを示せ．

12.15 $N(\mu,\sigma^2)$ からの大きさ n の標本による σ^2 に対する各不偏推定量の分散は少なくとも $2\sigma^4/n$ であることを示せ．

12.16 μ を既知として，$(x_1,\cdots,x_n)$ を正規分布 $N(\mu,\sigma^2)$ からの標本とする．このとき，$\dfrac{1}{n}\sqrt{\dfrac{\pi}{2}}\displaystyle\sum_{\xi=1}^{n}|x_\xi-\mu|$ は大きな n に対して漸近効率 $1/(\pi-2)$ を持つ σ の不偏推定量であることを示せ．

12.17 $(x_1,\cdots,x_n)$ は p.d.f.

$$\frac{1}{\pi[1+(x-\mu)^2]}$$

を持つコーシー分布からの標本である．このとき μ を推定する標本メディアンの漸近効率は大標本において $8/\pi^2$ であることを示せ．

12.18　$(x_1,\cdots,x_n)$ は p.d.f.

$$\frac{\lambda^{k+1}x^k e^{-\lambda x}}{\Gamma(k+1)}, \qquad x>0$$

を持つ分布からの標本である．ただし $\lambda>0$, k は既知の定数．このとき λ の最尤推定量 $\hat{\lambda}$ は $(k+1)/\bar{x}$ であることを示せ．またこの推定量には一致性はないが，不偏推定量であることを示せ．さらに大きな n に対するその漸近的分布は $N(\lambda,\lambda^2/[n(k+1)])$ であることも示せ．

12.19（続き）　$\sqrt{(k+1)}\log\hat{\lambda}$ は大きな n に対して漸近分布が分散 $\frac{1}{n}$ の正規分布となるような $\hat{\lambda}$ の関数であることを示せ．

12.20　$(x_1,\cdots,x_n)$ が 2 項分布 $Bi(1,p)$ からの標本であるとき，p に対する最尤推定量 $\hat{p}$ は $\bar{x}$ で，この大標本に対する漸近分布は $N(p,[p(1-p)]/n)$ であることを示せ．また $\sin^{-1}(2\hat{p}-1)$ は大標本の場合，分散 $\frac{1}{n}$ を持つ正規分布を漸近分布として持つことを示せ．

12.21　$\bar{x}$ は 2 項分布 $Bi(1,p)$ からの大きさ n の標本平均である．このとき，大きな n に対する漸近的最短 $100\gamma\%$ 信頼区間は次式を満足する p の 2 つの値によって与えられることを示せ．

$$\frac{(\bar{x}-p)\sqrt{n}}{\sqrt{p(1-p)}}=\pm y_\gamma.$$

ただし $p(-y_\gamma<u<+y_\gamma)=\gamma$, u は分布 $N(0,1)$ を持つ確率変数である．

12.22　$\bar{x}$ はポアソン分布 $Po(\mu)$ から大きさ n の標本の標本平均である．このとき，大きな n に対する μ の漸近的最短 $100\gamma\%$ 信頼区間は，次式を満足する μ の 2 つの値によって与えられることを示せ．

$$\frac{(\bar{x}-\mu)\sqrt{n}}{\sqrt{\mu}}=\pm y_\gamma$$

ただし y_γ は前問に定義されている．

12.23　$(x_{1\xi},\cdots,x_{k\xi},\ \xi=1,\cdots,n)$ は多項分布 $M(1;p_1,\cdots,p_k)$ からの大きさ n の標本である．このとき $(p_1,\cdots,p_k)$ に対する最尤推定量は $(\bar{x}_1,\cdots,\bar{x}_k)$ となることを示せ．ただし $\bar{x}_i=\frac{1}{n}\sum_{\xi=1}^{n}x_{i\xi}$ である．この推定量の共分散行列は

$$\left\|\frac{1}{n}(\delta_{ij}p_i-p_ip_j)\right\|$$

であることを示せ．ただし

$$\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j.\end{cases}$$

12.24　$(x_{1\xi},\cdots,x_{k\xi},\ \xi=1,\cdots,n)$ は k 次元正規分布 $N(\{\mu_i\},\|\sigma_{ij}\|)$ からの大きさ n の標本である．このとき，ベクトル $(\mu_1,\cdots,\mu_k)$ の最尤推定量は標本成分の平均ベクトル $(\bar{x}_1,\cdots,\bar{x}_k)$ であることを示せ．また行列 $\|\sigma_{ij}\|$ の最尤推定量は

$$\left\|\frac{n-1}{n}s_{ij}\right\|$$

であることを示せ．ただし

$$s_{ij}=\frac{1}{n-1}\sum_{\xi=1}^{n}(x_{i\xi}-\bar{x}_i)(\bar{x}_{j\xi}-\bar{x}_j).$$

12.25　$(y_1,\cdots,y_n)$ はそれぞれ正規分布

$$N(\beta_1x_{11}+\cdots+\beta_kx_{k1},\sigma^2),\cdots,N(\beta_1x_{1n}+\cdots+\beta_kx_{kn},\sigma^2)$$

を持つ独立な確率変数である．ただし $x_{11},\cdots,x_{k1},\cdots,x_{1n},\cdots,x_{kn}$ は定数である．

$$a_{pq}=\sum_{\xi=1}^{n}x_{p\xi}x_{q\xi}, \qquad p,q=1,\cdots,k$$

$$a_{p0}=\sum_{\xi=1}^{n}x_{p\xi}y_\xi, \qquad p=1,\cdots,k$$

とする．このとき $(\beta_1,\cdots,\beta_k)$ の最尤推定量は

$$\left(\sum_{p=1}^{k}a_{p0}a^{p1},\cdots,\sum_{p=1}^{k}a_{p0}a^{pk}\right)$$

であることを示せ．ただし $\|a^{pq}\|=\|a_{pq}\|^{-1}$, $p,q=1,\cdots,k$ で，$\|a_{pq}\|$ は正則であると仮定する．また σ^2 の最尤推定量は

$$\frac{1}{n}\cdot\frac{|a_{p_0q_0}|}{|a_{pq}|}, \qquad p_0,q_0=0,1,\cdots,k, \quad p,q=1,\cdots,k$$

であることを示せ．ただし $a_{00}=\sum_{\xi=1}^{n}y_\xi^2$ とする．

12.26　$(x_1,\cdots,x_n)$ は p.d.f. $f_n(x_1,\cdots,x_n;\theta_0)$ を持ち，$\tilde{\theta}$ は θ_0 の不偏推定量とする．このとき任意の $\delta>0$ に対して

$$\sigma^2(\tilde{\theta}|\theta_0)\geqq 1/[\mathscr{E}(g^2|\theta_0)]$$

となることを示せ．ただし

$$g=\frac{f_n(x_1,\cdots,x_n;\theta_0+\delta)-f_n(x_1,\cdots,x_n;\theta_0)}{\delta f_n(x_1,\cdots,x_n;\theta_0)}.$$

したがって

$$\sigma^2(\tilde{\theta}|\theta_0)\geqq 1\Big/\Big[\inf_{\delta}\mathscr{E}(g^2|\theta_0)\Big]$$

を示せ．この結果は Chapman と Robbins (1951) による．

12.27　$(x_1,\cdots,x_n)$ は p.d.f. $f_n(x_1,\cdots,x_n;\theta)$ を持つ確率変数とする．しかも f_n は (12.2.25) を満足する，すなわち θ の十分統計量 $\tilde{\theta}$ が存在するとする．$f_n(x_1,\cdots,x_n;$

θ) と $v(\tilde{\theta}\,;\theta)$ は $R_n \times \Omega$ の（確率 0 の点をのぞいた）各点でそれぞれ 2 次偏導関数 $\partial^2 f_n/(\partial x_\xi \partial\theta)$ と $\partial^2 v/(\partial x_\xi \partial\theta)$ を持つとする．ただし R_n は標本空間で，Ω はパラメータ空間（実線上の開区間）である．このとき $f_n(x_1,\cdots,x_n\,;\theta)$ は $\exp[K_1(\theta)g(\tilde{\theta})+K_2(\theta)+h_n(x_1,\cdots,x_n)]$ なる形をとることを示せ．ただし $K_1(\theta)$, $K_2(\theta)$ は $(x_1,\cdots,x_n)$ に依存せず，$h_n(x_1,\cdots,x_n)$ は θ に依存しない．$g(\tilde{\theta})$ は $\tilde{\theta}$ に依存し，θ には依存しない．[Koopman (1936) と Pitman (1936)].

12.28（続き）**平均値の最小分散推定量としての十分推定量の関数** [Rao (1949)]

$u(g(\tilde{\theta}))$を平均値が $q(\theta)$ である $\tilde{\theta}$ の関数とする．また $v(g(\tilde{\theta}))$ を平均値が $q(\theta)$ の別の関数とする．このとき $w(g(\tilde{\theta}))=u(g(\tilde{\theta}))-v(g(\tilde{\theta}))$ ならば，すべての $\theta\in\Omega$ に対して

$$\int_{R_n} w(g)\exp[K_1(\theta)g+K_1(\theta)+h_n]\,dx_1\cdots dx_n=0$$

となる．この式が積分記号の下で繰り返し微分可能で，$w(g)$ がテーラー展開可能ならば，すべての $\theta\in\Omega$ に対して $\mathscr{E}(w(g))^2=0$ となることを示せ．これからすべての $\theta\in\Omega$ に対して確率 1 で $w(g)=0$ となることがわかる．よって $u(g(\tilde{\theta}))$ は高々確率 0 の集合の上で $v(g(\tilde{\theta}))$ と異なり，この仮定の下では $u(g(\tilde{\theta}))$ は $g(\theta)$ の一意の不偏推定量である．したがって **12.2.3** より，$u(g(\tilde{\theta}))$ はこの平均値の推定量として最小分散を持つ．

第13章　パラメトリック仮説検定

13.1 序論と定義

本章では主として Neyman と Pearson (1928, 1933) によって，最初に構築されたパラメトリック仮説検定論の基礎を取り上げる．さらに本理論と大標本における最尤推定量に関する理論との関係を述べよう．

統計的仮説検定の概念はこの数年間に発展し一般化されてきた．しかし，ここでその展開を詳細に述べる意図はない．この詳細に興味を持たれた読者は Lehmann (1959) を参照されたい．

この節では有限標本に関する基本概念および定理を若干導入し，あとの節で大標本に対する漸近理論を取り扱おう．大標本に対する漸近理論は，特に第 12 章で述べた大標本に対するパラメトリック推定理論と関係が深い．

$(x_1,\cdots,x_n)$ は c.d.f. $F_n(x_1,\cdots,x_n\,;\theta)$ を持つ n 次元確率変数とし，θ は r 次元のパラメータでその空間を Ω とする．ただし，Ω はここではユークリッド空間 R_r における開領域とする．$(x_1,\cdots,x_n)$ の最も重要な場合は，$(x_1,\cdots,x_n)$ が c.d.f. $F(x\,;\theta)$ からの無作為標本のとき，すなわち $F_n(x_1,\cdots,x_n\,;\theta)=\prod_{\xi=1}^{n}F(x_\xi\,;\theta)$ のときである．本章ではほとんど無作為標本の場合を取り扱う．

ω を Ω の部分集合とする．θ_0 が θ の母集団における真値であれば，$\theta_0\in\Omega-\omega$ に対して $\theta_0\in\omega$ であるという統計的仮説 $\mathscr{H}$ を設定することになるだろう．この仮説を簡単に $\mathscr{H}(\omega\,;\Omega)$ と書く．$\theta_0\in\omega$ であれば仮説は真であるといい，$\theta_0\in\Omega-\omega$ ならば偽であるという．$\theta_0\in\omega$ なる仮説は**帰無仮説**ともいわれる．誤解のないかぎり，0 を落して θ_0 を θ で表わす．ω が 1 点 θ だけの場合，仮説 $\mathscr{H}$ を**単純仮説**といい，その他の場合は**複合仮説**という．$(x_1,\cdots,x_n)$ に基づいて $\mathscr{H}$ を採択するか棄却するかを決めるわけ

である．正確に述べると次のようになる．W_n を θ に依存しない標本空間 R_n における集合として，$(x_1, \cdots, x_n) \in W_n$ ならば $\mathscr{H}$ を棄却し，さもなければ採択する．ここで 2 種類の誤りが認識される．**第 1 種の誤り**は $\mathscr{H}$ が真であるのに $\mathscr{H}$ を棄却する誤りであり，**第 2 種の誤り**は $\mathscr{H}$ が偽であるのに $\mathscr{H}$ を採択する誤りである．W_n の選択および $\mathscr{H}$ を棄却するための基準になる W_n の中の $(x_1, \cdots, x_n)$ の生起を**統計的検定**という．または，簡単に $\mathscr{H}$ の**検定**という．W_n は検定の**棄却集合**または**棄却域**という．誤解がなければ簡単に W_n を $\mathscr{H}$ の**検定**と呼ぶ．

(13.1.1)*)
$$P(W_n|\theta) = \int_{W_n} dF_n$$

は検定 W_n の**検出力**といい，区間 $[0,1]$ 上の値をとる θ の関数である．θ の関数 $P(\overline{W}_n|\theta) = 1 - P(W_n|\theta)$ を検定 W_n の**検査特性関数**という．第 1 種および第 2 種の誤りをおかす確率はそれぞれ

(13.1.2)
$$P(W_n|\theta), \qquad \theta \in \omega$$

(13.1.3)
$$1 - P(W_n|\theta), \qquad \theta \in \Omega - \omega$$

である．この 2 つの確率は第 1 種および第 2 種の誤りの**危険**ともいわれる．

検定が満足のいくものであるためには，もちろんある基準を満たさなければならない．(13.1.2)，(13.1.3) で与えられる第 1 種および第 2 種の誤りに対する確率がともに 0 であれば，検定 W_n は理想的である．しかし，明らかにこの要求は一般的には望むべくもない．そこで少しゆるい条件で満足しなければならない．検定 W_n にとって望ましい性質は不偏性である．すなわち，与えられた有意水準 α $(0 < \alpha < 1)$ に対して条件

(13.1.4)
$$P(W_n|\theta) \leqslant \alpha, \qquad \theta \in \omega$$

(13.1.5)
$$P(W_n|\theta) > \alpha, \qquad \theta \in \Omega - \omega$$

を満足することである．これは次の事柄を意味している．第 1 種の誤りはその確率が α を越えなく，同時に，$\mathscr{H}$ が真でないとき（すなわち $\theta \in \Omega - \omega$ のとき），この仮説を棄却する確率が α 以上であることを保障している．検定 W_n がある $\theta \in \omega$ に対して $P(W_n|\theta) = \alpha$ であれば，W_n は θ に対する大きさ α の検定であるといい，$W_{n,\alpha}$ と表わす．誤解がなければ，与えられた α と標本の大きさ n に対する検定を今後，W_α を用いて表わそう．不偏でない検定を**偏り**があるという．

W_α がすべての $\theta \in \omega$ に対して，(13.1.4) の等号が成り立つような検定ならば，W_n

*) 集合 W_n の特性関数 φ_{W_n} を W_n の点で $\varphi_{W_n} = 1$，その他で $\varphi_{W_n} = 0$ として，$P(W_n|\theta) = \mathscr{E}(\varphi_{W_n}|\theta)$ を導入することもある．

は標本空間に相似，あるいは単に W_α は**相似棄却域**であるという．

もし

$$\lim_{n\to\infty} P(W_\alpha|\theta) = 1, \qquad \theta \in \Omega - \omega$$

ならば，W_α を仮説 $\mathscr{H}(\omega\,;\Omega)$ に対する大きさ α の**一致検定**という．この用語は Wald と Wolfowitz (1940) によって導入された．

W_α，W_α^* が次のような与えられた α に対する $\mathscr{H}(\omega\,;\Omega)$ の 2 つの検定であれば

(13.1.6)
$$P(W_\alpha^*|\theta) \leqslant P(W_\alpha|\theta), \qquad \theta \in \omega \text{ のとき}$$

(13.1.7)
$$P(W_\alpha^*|\theta) > P(W_\alpha|\theta), \qquad \theta \in \Omega - \omega \text{ のとき}$$

検定 W_α^* は検定 W_α よりも一様に強力であるという．すなわち，W_α^* は W より良好であると考えられる．R_n における任意の（ボレル）集合 W_α に対して (13.1.6) と (13.1.7) を満たすとき，検定 W_α^* は $\mathscr{H}$ に対する**一様最強力検定**であるという．

例題　上で述べた概念を重要な例題を用いて説明しよう．$(x_1, \cdots, x_n)$ を正規分布 $N(\mu, \sigma^2)$ を持つ母集団からの標本とする．複合仮説 $\mathscr{H}(\omega\,;\Omega)$ を考えよう．ただし，Ω は $\sigma^2 > 0$ なる点 (μ, σ^2) の集合である．（すなわち，ユークリッド半平面である）．一方 ω は $\mu = \mu_0$ なる Ω の部分集合である（すなわち，Ω における直線 $\mu = \mu_0$ である）．**8.4.3** より，$\mathscr{H}$ が真であれば，$\sqrt{n}(\bar{x} - \mu_0)/s$ はスチューデント分布 $S(n-1)$ を持つことがわかる．W_α を次式を満たす R_n の点集合とする．

$$\left|\frac{(\bar{x} - \mu_0)\sqrt{n}}{s}\right| > t_\alpha.$$

ただし

$$1 - \int_{-t_\alpha}^{+t_\alpha} f_{n-1}(t)\,dt = \alpha.$$

ここに，$f_{n-1}(t)$ は (7.8.4) で与えられるスチューデント分布 $S(n-1)$ の p.d.f. である．このとき

$$P(W_\alpha|(\mu, \sigma^2) \in \omega) = P(|t| > t_\alpha) = \alpha$$

である．よってこの W_α は標本空間に相似な $\mathscr{H}(\omega\,;\Omega)$ に対する大きさ α の検定である．さらに，$\Omega - \omega$ における任意の点 (μ_1, σ_1^2) に対して

$$P(W_\alpha|(\mu_1, \sigma_1^2)) = P\left(\left|\frac{(\bar{x} - \mu_0)\sqrt{n}}{s}\right| > t_\alpha|(\mu_1, \sigma_1^2)\right) = P\left(\left|t + \frac{\delta}{\sqrt{u}}\right| > t_\alpha\right).$$

ただし $\delta = (\mu_1 - \mu_0)\sqrt{n(n-1)}/\sigma_1$ で (t, u) は半平面 $u \geqslant 0$ 上で p.e.

$$\frac{\left(\dfrac{u}{2}\right)^{\frac{1}{2}(n-3)}}{2\sqrt{\pi(n-1)}\,\Gamma\left(\dfrac{n-1}{2}\right)} e^{-\frac{1}{2}(1+t^2/(n-1))u}\,dt\,du$$

を持つ確率変数である．この p.e. を $f(t,u)\,dt\,du$ で表わす．$(0,\infty)$ における u のすべての値に対して

$$\int_{-t_\alpha-\frac{\delta}{\sqrt{u}}}^{+t_\alpha-\frac{\delta}{\sqrt{u}}} f(t,u)\,dt < \int_{-t_\alpha}^{+t_\alpha} f(t,u)\,dt$$

となることに注意すると

$$P\left(\left|t+\frac{\delta}{\sqrt{u}}\right|>t_\alpha\right) > P(|t|>t_\alpha)=\alpha, \qquad \delta \doteqdot 0$$

が与えられる．もちろん，これは W_α が $\mathscr{H}$ に対する不偏検定であることを意味している．

不都合なことに，不偏かつ一様最強力な検定は種々の特殊な場合にのみ存在する．単純仮説の場合は，$\Omega-\omega$ が 1 点からなっているときにはかなり一般的な解が存在し，$\Omega-\omega$ が 1 点より多い集合の場合は，ある種の条件の下で解が存在する．単純仮説については 13.2 節で考察しよう．複合仮説の場合には，解は特殊な条件のときのみしか存在しない．特に W_α が標本空間に相似なとき存在する．しかし最尤推定量が存在して，この推定量が漸近的に正規分布に従うような母集団分布の場合においては，大標本に対する漸近的な意味で，ある複合仮説の一様最強力不偏検定が存在する．この問題は，13.3，13.4 節で取り上げられる．

歴史的注意　大量生産品のロットの採択，棄却に対する抜取検査論を展開する過程において，ベル電話研究所のドッジとロミッグは 1925 年ごろ生産者危険および消費者危険という概念を導入している．最初の結果はその 4 年後に発表されている [Dodge と Romig (1929, 1959) を見よ]．これらの概念は実は (13.1.2)，(13.1.3) で与えられた第 1 種と第 2 種の誤りの危険の先駆をなしている．厳密に表わそう．θ を N 個の製品からなるロットにおける不良品の割合，t をロットから抜き取られた n 個の製品からなる標本における不良品の割合としよう．このとき，ロットにおける θ の与えられた値に対して，t は c.d.f. $F_n(t\,;\theta)$ を持つ確率変数になる．t の標本空間は区間 $[0,1]$ の $0,\dfrac{1}{n},\cdots,\dfrac{r}{n}\ (r\leqslant n)$ であり，θ のパラメータ空間は区間 $[0,1]$ 上の点 $0,\dfrac{1}{N},\cdots,\dfrac{N}{N}$ である，θ の値 θ_0 が与えられると，ロットに対する不採択不良率は $\theta>\theta_0$ なる値で，この場合，$\omega=(\theta_0,1)$ になる．一方，採択不良率（優先採択領域）は $\theta\leqslant\theta_0$ なる値で，この場合，$\Omega-\omega=(0,\theta_0)$ になる．さて，t_α および棄却域 $W_\alpha=[t_\alpha,1]$ を，$t\in W_\alpha$ ならロットを棄却し，しかも $P(t\in W_\alpha|\theta\in\omega)\leqslant\alpha$ となるように選ぶ．$P(t\in W_\alpha|\theta\in\omega)$ は**生産者危険**である．すなわちロットの不良率が θ のとき，検査官がロットを棄却する確率である．$\theta=\theta_0$ のとき（近似的に）α を定めておいて，生産者危険は α を越えないように制御される．実際には α の値として 0.05 あるいは 0.10 を一般に用いる，$P(t\in\overline{W}_\alpha|\theta\in\Omega-\omega)=\beta$ は**消費者危険**である．すなわち，ロットが不良率 θ であるのに，検査官がこのロットを採択する確率である．事実，ロットの不良率が θ_0 を越える任意の特定の値ならば，N が十分大きいとき，n を十分大きく選べば消費者危険を適当に小さくするようにできる．通常，θ_0 より少し大きな θ のある与えられた値 θ_1 に対して，消費者危険 β が与えられた値（0.05 または 0.10 がよく用いられる）を持つように標本の大きさ n を選ぶ．θ の不採択値の区間 $(\theta_0,1)$ は 2 つの集合に分けられる．すなわち，**無関心領域** (θ_0,θ_1) および**優先棄却領域** $(\theta_1,1)$ である．任意に与えられた不良率 θ を持つロットからの標本の大きさ n に対して $P(t\in\overline{W}_\alpha|\theta)$ を計算することができる．これは不良率が θ のとき，ロットを採択する確率である．この θ の関数を $L(\theta)$ で表わす．$L(\theta)$ のグラフは標本の大きさ n と棄却不良率 t_α とで表わされる検査方式の**検査特性曲線**という．この関数は $L(\theta_0)=1-\alpha$，$L(\theta_1)=\beta$ を満足する．

Dodge と Romig (1929) が導入したもう 1 つの重要な概念は**平均検出品質限界**である．これは次のようなロットの不良率の平均値の最大値（θ の値全体にわたる）として定義される．すなわち棄却されるロットの不良品が取り除かれて，採択されるようになるロットの不良率の平均値全体にわたる最大値である．この検査は，もちろん，非破壊検査としている．すなわち，不良品であるかどうかは，その製品を壊さないで検査できるものとする．

第 1 種と第 2 種の誤りの基本概念は刑事法の施行上，根本的な役割を果たしてきている．第 1 種，第 2 種の誤りの危険はそれぞれ無実の人間に有罪の判決を下す危険，罪人を放免する危険に相当する．

13.2　単純仮説検定

(a)　対立する仮説が 2 つの場合

$\mathscr{H}$ が，ω における 1 点 θ_0 と，$\Omega-\omega$ における 1 点 θ_1 をとることで表わされる単純仮説の場合，$(x_1,\cdots,x_n)$ が θ_0 と θ_1 に対する p.d.f. $f_n(x_1,\cdots,x_n\,;\theta)$ を持つならば，次に示す条件の下では，$\mathscr{H}(\theta_0\,;\theta_0\cup\theta_1)$ に対する不偏最強力検定が存在する．この問題は Neyman と Pearson (1933) により最初に研究された．次の定理はこの 2 人によるもので，$\mathscr{H}$ に対する不偏最強力検定を与えている．

13.2.1　$(x_1,\cdots,x_n)$ は確率変数で p.d.f. $f_n(x_1,\cdots,x_n\,;\theta)$ を持ち，パラメータ空間 Ω は 2 つの相異なる点（簡単のために θ_0，θ_1 で表わす）からなるものと仮定する．与えられた $c_\alpha>0$ に対して，W_α を次の条件を満たす標本空間

R_n における集合とする.

(13.2.1) $$f_n(x_1, \cdots, x_n ; \theta_1) \geqslant c_\alpha f_n(x_1, \cdots, x_n ; \theta_0).$$

ただし

$$P(W_\alpha | \theta_0) = \alpha.$$

もし W_α^* が $P(W_\alpha^* | \theta_0) = \alpha$ なる R_n における任意の他の集合であれば

(13.2.2) $$P(W_\alpha | \theta_1) \geqslant P(W_\alpha^* | \theta_1).$$

すなわち, W_α は不偏で, かつ他の任意の大きさ α の検定 W_α^* よりも $\mathscr{H}(\theta_0 ; \theta_0 \cup \theta_1)$ の検定としては強力である.

13.2.1 を証明しよう. まず W_α は W_α^* よりも強力な検定であることを示そう. 与えられた $c_\alpha > 0$ に対して, W_α が (13.2.1) を満たす $(x_1, \cdots, x_n)$ 空間における事象であり, $\theta = \theta_0$ に対する W_α の確率が α であることに注目する. すなわち, W_α は $\mathscr{H}(\theta_0 ; \theta_0 \cup \theta_1)$ に対する大きさ α の検定である.

いま, W_α^* が大きさ α の $\mathscr{H}$ に対する他の任意の検定であるとしよう. すなわち

(13.2.3) $$P(W_\alpha^* | \theta_0) = \alpha.$$

よって

(13.2.4) $$P(W_\alpha - (W_\alpha \cap W_\alpha^*) | \theta_0) = P(W_\alpha^* - (W_\alpha \cap W_\alpha^*) | \theta_0).$$

さて W_α に含まれる標本空間 R_n の任意の点に対して, $f_n(x_1, \cdots, x_n ; \theta_1) \geqslant c_\alpha f_n(x_1, \cdots, x_n ; \theta_0)$ である. よって

(13.2.5) $$P(W_\alpha - (W_\alpha \cap W_\alpha^*) | \theta_1) \geqslant c_\alpha P(W_\alpha - (W_\alpha \cap W_\alpha^*) | \theta_0).$$

W_α に含まれない点に対しては, $f_n(x_1, \cdots, x_n ; \theta_1) < c_\alpha f_n(x_1, \cdots, x_n ; \theta_0)$ であるので

(13.2.6) $$c_\alpha P(W_\alpha^* - (W_\alpha \cap W_\alpha^*) | \theta_0) \geqslant P(W_\alpha^* - (W_\alpha \cap W_\alpha^*) | \theta_1).$$

(13.2.4) を (13.2.5) と (13.2.6) に用いれば

(13.2.7) $$P(W_\alpha - (W_\alpha \cap W_\alpha^*) | \theta_1) \geqslant P(W_\alpha^* - (W_\alpha \cap W_\alpha^*) | \theta_1).$$

さて, $W_\alpha - (W_\alpha \cap W_\alpha^*)$ と $W_\alpha \cap W_\alpha^*$ は互いに素な集合であるので, $W_\alpha^* - (W_\alpha \cap W_\alpha^*)$ と $W_\alpha \cap W_\alpha^*$ も互いに素となる. そこで $P(W_\alpha \cap W_\alpha^* | \theta_1)$ を (13.2.7) の両辺に加えれば (13.2.2) が得られる. すなわち

$$P(W_\alpha | \theta_1) \geqslant P(W_\alpha^* | \theta_1).$$

これは, W_α が $\mathscr{H}$ の検定としては W_α^* より強力であることを示している.

次に W_α の不偏性を証明する.

いま, (13.2.1) で $c_\alpha \geqslant 1$ ならば, W_α の不偏性は明白である. $c_\alpha < 1$ であれば,

(13.2.8) $$f_n(x_1, \cdots, x_n ; \theta_1) \geqslant f_n(x_1, \cdots, x_n ; \theta_0), \qquad (x_1, \cdots, x_n) \in W_\alpha' \text{ のとき}$$

となる W_α の部分集合 W_α' が存在する. そうでなければ, R_n のいたるところで

(13.2.9) $$0 \leqslant f_n(x_1, \cdots, x_n ; \theta_1) \leqslant f_n(x_1, \cdots, x_n ; \theta_0)$$

となるが, これは

$$\int_{R_n} f_n(x_1, \cdots, x_n ; \theta_0)\, dx_1 \cdots dx_n = \int_{R_n} f_n(x_1, \cdots, x_n ; \theta_1)\, dx_1 \cdots dx_n = 1$$

に矛盾する. R_n が全標本空間であれば

(13.2.10) $$\begin{aligned} P(W_\alpha' | \theta_1) - P(W_\alpha' | \theta_0) &= P(R_n - W_\alpha' | \theta_0) - P(R_n - W_\alpha' | \theta_1) \\ &\geqslant P(W_\alpha - W_\alpha' | \theta_0) - P(W_\alpha - W_\alpha' | \theta_1) \end{aligned}$$

は明らかである. すなわち

(13.2.11) $$P(W_\alpha | \theta_1) \geqslant P(W_\alpha | \theta_0) = \alpha.$$

よって W_α は不偏である.

13.2.1 と同様な定理は, $(x_1, \cdots, x_n)$ が p.f. $p_n(x_1, \cdots, x_n ; \theta)$ を持つ離散型 n 次元確率変数の場合にも成り立つ. このとき p.f. $p_n(x_1, \cdots, x_n ; \theta_0)$ および $p_n(x_1, \cdots, x_n ; \theta_1)$ を持つ分布は統計的仮説 $\mathscr{H}(\theta_0 ; \theta_0 \cup \theta_1)$ に含まれる相対立する分布である.

最後に, $(x_1, \cdots, x_n)$ が p.d.f. $f(x ; \theta)$ を持つ分布からの確率変数である場合には, **13.2.1** の $f_n(x_1, \cdots, x_n ; \theta)$ を $\prod_{\xi=1}^{n} f(x_\xi ; \theta)$ で置き換えればよいことに注意しよう. 同様な事柄が $(x_1, \cdots, x_n)$ が p.f $p(x ; \theta)$ を持つ離散型分布からの標本にも成り立つ.

(b) 相似な棄却域の場合

ある特定の p.d.f. $f_n(x_1, \cdots, x_n ; \theta)$ については, c_α のすべての正の値と対立仮説 $\theta = \theta_1$ に対して (13.2.1) で定義された集合族は, (13.2.2) において $\theta_1 \neq \theta_0$ となる Ω のすべての θ_1 の値に対する集合族と同じになる. この場合, W_α は仮説 $\mathscr{H}(\theta_0 ; \Omega)$ に対する大きさ α の一様不偏最強検定である.

特に p.d.f. $v(\tilde{\theta} ; \theta)$ を持つ θ に対する十分推定量 $\tilde{\theta}$ が存在して, c_α のすべての正の値に対して $v(\tilde{\theta} ; \theta_1) \geqslant c_\alpha v(\tilde{\theta} ; \theta_0)$ で定義される R_n の集合族が, $\theta_1 \neq \theta_0$ となる Ω のすべての θ_1 の値から得られるような集合族と同じになれば, 与えられた α に対してこの族から $\mathscr{H}(\theta_0 ; \Omega)$ に対する一様不偏最強検定 W_α を求めることができる. これは明らかに, 十分推定量 $\tilde{\theta}$ により決められる. しかし, 1 次元パラメータ θ の場合以外は, こ

の方法ではほとんど望みがない．

ここで，このような検定の存在を確かめるために例を示そう．

例題　$(x_1, \cdots, x_n)$ を $N(\mu, 1)$ からの標本とする．μ のパラメータ空間を実数軸 R_1 とする．このとき，仮説 $\mathscr{H}(\mu_0\,;\,\mu_0 \cup \mu_1)$ を考えよう．(13.2.1) を用いて，変形すれば，W_α は

$$e^{n(\mu_1-\mu_0)\bar{x}} \geqslant c'_\alpha$$

を満たす標本空間 R_n における集合であることがわかる．ただし，$c'_\alpha > 0$．このように，W_α は μ に対する十分推定量 $\bar{x}$ によって定まる．$(x_1, \cdots, x_n)$ が $N(\mu, 1)$ からの標本であれば，$\bar{x}$ が分布 $N\left(\mu, \dfrac{1}{n}\right)$ を持つことに注意すれば，次の事柄は明らかである．

(i) $\mu_1 > \mu_0$ ならば，W_α は $\bar{x} > z_\alpha$ なる R_n の集合である．ただし，z_α は方程式

$$1 - \Phi(\sqrt{n}(z_\alpha - \mu_0)) = \alpha$$

で定義され，$\Phi(x)$ は $N(0, 1)$ の c.d.f. である．

(ii) $\mu_1 < \mu_0$ ならば，W_α は $\bar{x} < z'_\alpha$ となる集合である．ただし

$$\Phi(\sqrt{n}(\mu_0 - z'_\alpha)) = \alpha.$$

(i) の場合，$\bar{x} > z_\alpha$ で定義される棄却域を持つ検定は，$\mathscr{H}(\mu_0\,;\,\mu \geqslant \mu_0)$，すなわち，任意の対立仮説 $\mu = \mu_1$ に対して $\mu = \mu_0$ であるという仮説に対する一様最強力検定である．ただし $\mu_1 > \mu_0$．また，

$$1 - \Phi(\sqrt{n}(z_\alpha - \mu_1)) > 1 - \Phi(\sqrt{n}(z_\alpha - \mu_0)), \qquad \mu_1 > \mu_0 \text{ のとき}$$

だから，Ω が $\mu \geqslant \mu_0$ なるすべての μ の値からなれば，この検定は明らかに不偏である．

同様な事柄が (ii) の場合にもいえる．

(c) $\Omega - \omega$ において 1 点を持つ複合仮説の単純仮説へのワルドの還元

複合仮説 $\mathscr{H}(\omega\,;\,\omega \cup \theta_1)$，すなわち $\Omega - \omega$ が 1 点 θ_1 を持つ場合に関して，Wald (1939) は $\mathscr{H}$ の検定方式の構成問題を次のように取りあげている．これは，本質的には $\theta \in \omega$ に対する p.d.f. $f_n(x_1, \cdots, x_n\,;\,\theta)$ の族を，θ に関して，ω 上で定義された先験的 c.d.f. $G(\theta)$ で平均して得られる p.d.f. $f_G(x_1, \cdots, x_n)$ で置き換える方法である．すなわち

$$f_G(x_1, \cdots, x_n) = \int_\omega f_n(x_1, \cdots, x_n\,;\,\theta)dG(\theta) \tag{13.2.12}$$

とする．これは p.d.f. である．ワルドは，母集団 p.d.f. は $f_G(x_1, \cdots, x_n)$ であるという仮説 $\mathscr{H}_G$ を，母集団 p.d.f. は $f_n(x_1, \cdots, x_n\,;\,\theta_1)$ であるという対立仮説に対して検定する問題を考察している．

$f_G(x_1, \cdots, x_n)$ が **13.2.1** における $f_n(x_1, \cdots, x_n\,;\,\theta)$ と同じ性質すべてを持っていると

仮定すれば，**13.2.1** は仮説 $\mathscr{H}_G$ に対する大きさ α の最強力検定 $W_{G\alpha}$ を定めることに応用できる．Wald (1939) の定理は次の通りである．

13.2.2　仮説 $\mathscr{H}_G$ に対する大きさ α の最強力検定 $W_{G\alpha}$ が，また $\mathscr{H}(\omega\,;\,\omega \cup \theta_1)$ の大きさ α の検定となるような c.d.f. $G(\theta)$ $(\theta \in \omega)$ が存在すれば

$$\int_{\omega_0} dG(\theta) = 0 \tag{13.2.13}$$

である．ただし ω_0 は

$$\int_{W_{G\alpha}} f_n(x_1, \cdots, x_n\,;\,\theta)\,dx_1 \cdots dx_n = \alpha, \tag{13.2.14}$$

を満たす ω の部分集合である．さらに $W_{G\alpha}$ は仮説 $\mathscr{H}(\omega\,;\,\omega \cup \theta_1)$ に対する最強力検定である．

13.2.2 を証明しよう．与えられた α に対して，$W_{G\alpha}$ が $\mathscr{H}(\omega\,;\,\omega \cup \theta_1)$ に対する大きさ α の検定であれば

$$\int_{W_{G\alpha}} f_n(x_1, \cdots, x_n\,;\,\theta)\,dx_1 \cdots dx_n \leqslant \alpha, \qquad \theta \in \omega \tag{13.2.15}$$

となるはずであり，よって

$$\int_\omega \int_{W_{G\alpha}} f_n(x_1, \cdots, x_n\,;\,\theta)\,dx_1 \cdots dx_n\,dG(\theta) \leqslant \alpha \int_\omega dG(\theta). \tag{13.2.16}$$

積分の順序交換によって，

$$\int_{W_{G\alpha}} f_G(x_1, \cdots, x_n)\,dx_1 \cdots dx_n = \alpha$$

が成立するためには，すなわち，$W_{G\alpha}$ が大きさ α であるためには，(13.2.13) が成立しなければならないことがわかる．

さて，$W_{G\alpha}$ が $\mathscr{H}_G$ に対する最強力検定ならば，

$$\int_{W_{G\alpha}} f_n(x_1, \cdots, x_n\,;\,\theta_1)\,dx_1 \cdots dx_n > \int_{W'_{G\alpha}} f_n(x_1, \cdots, x_n\,;\,\theta_1)\,dx_1 \cdots dx_n \tag{13.2.17}$$

でなければならない．ただし $W'_{G\alpha}$ は $\mathscr{H}_G$ に対する大きさ α の他の任意の検定である．しかし (13.2.17) はまさしく，$W_{G\alpha}$ が $\mathscr{H}(\omega\,;\,\omega \cup \theta_1)$ に対する最強力検定になる条件そのものである．

$W_{G\alpha}$ が相似な集合であれば，すなわち，$\theta \in \omega$ に対して (13.2.15) の等号が成り立てば，(13.2.13) が正しいことは自明である．しかし，この例では，相似性は本質的に棄却域を定める問題を解くものである．

13.3 尤度比検定

有限標本に対して最適な特性を兼ね備えている複合仮説検定方式が，Neyman と Pearson (1928, 1933) による**尤度比原理**といわれる重要な原理により，種々の特定の問題に対して得られている．この検定方式は (13.2.2) で定義した検定を複合仮説の場合にごく自然な型で拡張したものである．

大標本においては尤度比検定は，漸近的に正規分布する最尤推定量が存在するための条件と同じ条件の下では最適な漸近性を持つ．

他方，有限標本に対する複合仮説検定の問題については，Lehmann (1950, 1959), Lehmann と Scheffé (1950), Lehmann と Stein (1948) などによりさらに一般的な方式が開発されている．

ここで詳細に考察する複合仮説検定は尤度比検定だけにしよう．特に n の大きな値に対する漸近的な性質を議論しよう．大きな n の値に対する尤度比検定は，第12章で考察したパラメトリック推定の大標本論と同様に，統計理論において基本的である．

検定される仮説が真のときの，大標本における複合仮説に対する尤度比検定の漸近分布は，Wilks (1938 a) によって最初に論じられた．Wald (1941 a, 1941 b) はその後，不偏性，検出力およびその他の概念の漸近的性質を含む尤度比検定を詳細に研究している．以下の数節で論じている結果はほとんどワルドの基本的な論文に含まれている結果を少し平易にしたものである．しかし，ここで用いられている方法はワルドの方法より，いくぶん単純化されている．

(a) 尤度比検定の定義

$(x_1, \cdots, x_n)$ は θ が r 次元であるような c.d.f. $F(x\,;\theta)$ からの標本とする．便宜上，x を 1 次元として取り扱うが，以下の結果は記号だけ変えれば k 次元に拡張できる．尤度素分は

(13.3.1)
$$dF_n(x_1, \cdots, x_n\,;\theta) = \prod_{\xi=1}^{n} dF(x_\xi\,;\theta)$$

である．さて，仮説

(13.3.2)
$$\mathscr{H}(\omega\,;\Omega)$$

を考えよう．また

(13.3.3)
$$(dF_n)_\omega = \sup_{\theta\in\omega} dF_n(x_1, \cdots, x_n\,;\theta)$$

(13.3.4)
$$(dF_n)_\Omega = \sup_{\theta\in\Omega} dF_n(x_1, \cdots, x_n\,;\theta)$$

とする．通常の適用例においては，ほとんど，$(dF_n)_\omega$ と $(dF_n)_\Omega$ は普通の微分演算で得られる $\theta\in\omega$ と $\theta\in\Omega$ に対する dF_n の最大値である．

仮説 $\mathscr{H}(\omega\,;\Omega)$，あるいは単に $\mathscr{H}$ の検定に対する**尤度比検定** $\lambda_{\mathscr{H}}$ は

(13.3.5)
$$\lambda_{\mathscr{H}} = \frac{(dF_n)_\omega}{(dF_n)_\Omega}$$

で定義される．$\lambda_{\mathscr{H}}$ の値は区間 $[0, 1]$ 上にある．仮説 $\mathscr{H}$ の検定に対する R_n における棄却域 W_α は $\lambda_{\mathscr{H}} < c_\alpha$ を満たす集合である．ここに c_α は定数で，

(13.3.6)
$$\int_{W_\alpha} dF_n \leqslant \alpha, \qquad \theta\in\omega$$

である．$\lambda_{\mathscr{H}} < c$ で定義される検定は，本質的には (13.2.2) で定義される検定の ω, $\Omega-\omega$ がおのおの 1 点以上からなっている場合への拡張である．

さらに $\lambda_{\mathscr{H}}$ は直感的に検定とわかる．これは，標本 $(x_1, \cdots, x_n)$ が与えられていて，θ のある値の“もっともらしさ”を他の値と比較するとき，尤度素分に大きな値を与える θ の値を直感的に選ぶことからもわかる．このように，集合 ω のすべての θ の値を探した尤度素分の値よりも，いくぶんでも大きな尤度素分の値がパラメータ空間 Ω 全体を探しても得られなければ，ω に属する θ の値を“一番もっともらしい”，すなわち仮説 $\mathscr{H}$ は真である，という命題が成り立っていると直感的に評価してもよかろう．大標本の場合に，この直感がかなり一般的な条件の下では厳密に成立することを見よう．

例題　$(x_1, \cdots, x_n)$ は $N(\mu, \sigma^2)$ からの標本としよう．つぎのように定義された仮説 $\mathscr{H}$ に対する尤度検定を求めたい．

Ω：　$-\infty < \mu < +\infty$, $\sigma^2 > 0$ なる半平面

ω：　$\mu = \mu_0$ なる Ω（半平面）の部分集合．

このとき，

$$dF_n = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}n} \exp\left[-\frac{1}{2\sigma^2}\sum_{\xi=1}^{n}(x_\xi - \mu)^2\right] dx_1 \cdots dx_n$$

を得る．Ω における (μ, σ^2) に関して dF_n を最大にし，また ω において (μ, σ^2) に関しても最大にする．そして (13.3.5) のように 2 つの最大値の比をとると $\mathscr{H}$ に対する尤度比として

$$\lambda_{\mathscr{H}} = \left(\frac{S_\Omega}{S_\omega}\right)^{\frac{1}{2}n}$$

が得られる．ただし，S_Ω は μ に関する平方和 $\sum_{\xi=1}^{n}(x_\xi-\mu)^2$ の最小値，すなわち

$$S_\Omega = \sum_{\xi=1}^{n}(x_\xi-\bar{x})^2.$$

他方，

$$S_\omega = \sum_{\xi=1}^{n}(x_\xi-\mu_0)^2.$$

$\lambda_{\mathscr{H}}$ は次のように書ける．

$$\lambda_{\mathscr{H}} = \left(1+\frac{t^2}{n-1}\right)^{-\frac{1}{2}n}, \qquad t=\sqrt{n}(\bar{x}-\mu_0)/s.$$

$\bar{x}$ と s^2 は (8.2.1)，(8.2.6) で定義された標本和と標本分散である．もちろん確率変数 t は (8.4.16) で定義された"スチューデント"比である．よって，尤度比検定の大きさ α の棄却域 W_α は，$\mu=\mu_0$ に対して

$$P(\lambda_{\mathscr{H}}<\lambda_\alpha)=\alpha$$

を満たす $\lambda_{\mathscr{H}}$ の標本空間の一部である．ここに

$$\lambda_\alpha = \left(1+\frac{t_\alpha^2}{n-1}\right)^{-\frac{1}{2}n}.$$

t_α は

$$\int_{-t_\alpha}^{+t_\alpha} f_{n-1}(t)\,dt = 1-\alpha$$

となるように選ばれ，$f_{n-1}(t)$ は (7.8.4) で定義された"スチューデント"分布 $S(n-1)$ の p.d.f. である．

したがって，仮説 $\mathscr{H}$ を検定する尤度比 $\lambda_{\mathscr{H}}$ は"スチューデント" t-検定と同値である．

(b) 正規回帰論における尤度比検定

尤度比検定の小標本への適用の 1 つに正規分布のパラメータに関する種々の仮説検定がある．前述の例題はその 1 つの例である．これらの応用例の多くは，尤度比の標本分布を大きさ n の有限標本に対して求めることができる．この尤度比検定はほとんど自明で，本章の終わりの問題をのぞいては，詳細な考察を加える必要はなかろう．

しかし，正規回帰論における回帰パラメータのいくつかが 0 の値を持つという仮説の検定に対しては，尤度比検定は考察するに価する．10.3 (c) 節に戻って，y_ξ，$\xi=1,\cdots,n$ は正規分布 $N(\beta_1x_{1\xi}+\cdots+\beta_kx_{k\xi},\sigma^2)$，$\xi=1,\cdots,n$ を持つ独立な確率変数であると仮定する．ただし，$(x_{i\xi},\ i=1,\cdots,k)$ は 1 次独立な固定されたベクトル（すなわち (10.3.8) の行列は正則）とする．$\mathscr{H}$ を Ω および ω が次のように定義されている仮説とする．

(13.3.7)
Ω：$-\infty<\beta_i<+\infty$，$i=1,\cdots,k$，$\sigma^2>0$ なる $(k+1)$ 次元半ユークリッド空間

ω：$\beta_{k'+1}=\cdots=\beta_k=0$，$k'<k$ なる Ω の部分集合

このように定義された仮説 $\mathscr{H}$ は正規回帰論の**一般線形仮説**といわれる．ここで取り扱う標本は条件つき確率変数 $(y_\xi|x_{1\xi},\cdots,x_{k\xi},\xi=1,\cdots,n)$ からなるものであり，パラメータ $\beta_1,\cdots,\beta_k$，σ^2 の尤度は

$$dF_n = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{\xi=1}^{n}(y_\xi-\beta_1x_{1\xi}-\cdots-\beta_kx_{k\xi})^2\right]dy_1\cdots dy_n \tag{13.3.8}$$

で与えられる．微分して $(dF_n)_\Omega$，$(dF_n)_\omega$ を求めると

$$\lambda_{\mathscr{H}} = \frac{(dF_n)_\Omega}{(dF_n)_\omega} = \left(\frac{S_\Omega}{S_\omega}\right)^{\frac{1}{2}n} \tag{13.3.9}$$

となる．ただし S_Ω は 2 乗和

$$\sum_{\xi=1}^{n}(y_\xi-\beta_1x_{1\xi}-\cdots-\beta_kx_{k\xi})^2 \tag{13.3.10}$$

の $\beta_1,\cdots,\beta_k$ に関する最小値で，S_ω は 2 乗和

$$\sum_{\xi=1}^{n}(y_\xi-\beta_1x_{1\xi}-\cdots-\beta_{k'}x_{k'\xi})^2 \tag{13.3.11}$$

の $\beta_1,\cdots,\beta_{k'}$ に関する最小値である．10.3 (b) 節を参照すると，S_Ω は (10.3.27) で与えられる S_1 と等しいことがわかる．すなわち

$$S_\Omega = \frac{|a_{i_0j_0}|}{|a_{ij}|}. \tag{13.3.12}$$

ここに a_{ij}，$i,j=1,\cdots,k$ および $a_{i_0j_0}$ は (10.3.8)，(10.3.11)，(10.3.28) で定義されている．もちろん $|a_{ij}|$ は行列 $\|a_{ij}\|$ の行列式で，$|a_{i_0j_0}|$ は (10.3.28) で定義されている．

同様にして

$$S_\omega = \frac{|a_{i_0'j_0'}|}{|a_{i'j'}|} \tag{13.3.13}$$

となる．ただし $|a_{i_0'j_0'}|$，$|a_{i'j'}|$ は $i',j'=1,\cdots,k'$ なる $|a_{i_0j_0}|$，$|a_{ij}|$ と同様である．

コクランの定理 8.4.4 に基づく議論より次のことがわかる．仮説 $\mathscr{H}$ が真のとき，S_Ω/σ^2

と $(S_\omega - S_\Omega)/\sigma^2$ はそれぞれカイ 2 乗分布 $C(n-k)$, $C(k-k')$ に従って独立に分布する．この詳細はあえてここで述べる必要はなかろう．ただ，S_1, S_2 が (10.3.30) で与えられたとき， S_1/σ^2 と S_2/σ^2 がそれぞれカイ 2 乗分布 $C(n-k)$, $C(k)$ に従って独立に分布するのとまったく同じである．

よって (13.3.9) における尤度比 $\lambda_{\mathscr{H}}$ は

(13.3.14)
$$\lambda_{\mathscr{H}} = \left(1+\frac{(k-k')}{(n-k)}\mathscr{F}\right)^{-\frac{1}{2}n}$$

と書ける．ここに

(13.3.15)
$$\mathscr{F} = \frac{(n-k)(S_\omega - S_\Omega)}{(k-k')S_\Omega}$$

は p.e. が (7.8.9) で定義されたスネディッカー分布 $S(k-k', n-k)$ を持つ．$\lambda_{\mathscr{H}}$ と $\mathscr{F}$ の間には 1 対 1 の対応があるので，$\mathscr{H}$ の尤度比検定は $\mathscr{F}$ 検定と同値である．$\lambda_{\mathscr{H}}$ の標本空間における棄却域 $\lambda_{\mathscr{H}} < \lambda_\alpha$ は（ただし λ_α は $\lambda_{\mathscr{H}}$ 分布の $100\alpha\%$ 点である），$\mathscr{F}$ の標本空間における棄却域 $\mathscr{F} > [(n-k)/(k-k')](\lambda_\alpha^{-2/n}-1)$ に対応している．この $\mathscr{F}$ 検定はモデル I 分散分析検定の最も一般化された形である．Daly (1940) はこの検定の不偏性を示している．

以上の考察より，正規回帰論における一般線形仮説に関する基本的な結果は次のようになる．

13.3.1　y_ξ, $\xi=1,\cdots,n$ を正規分布 $N(\beta_1x_{1\xi}+\cdots+\beta_kx_{k\xi},\sigma^2)$, $\xi=1,\cdots,n$ を持つ独立な確率変数とする．ただし，(10.3.8) で定義された行列 $\|a_{ij}\|$, $i,j=1,\cdots,k$ は正則行列とする．(13.3.7) で示される仮説 $\mathscr{H}$ の検定に対する尤度比 $\lambda_{\mathscr{H}}$ は (13.3.14) で与えられる．ただし，$\mathscr{F}$ は (13.3.15) で与えられている．さらに，$\mathscr{H}$ が真であれば，$[(n-k)/(k-k')](\lambda_{\mathscr{H}}^{-2/n}-1)$ はスネディッカー分布 $S(k-k',n-k)$ を持つ．

一般線形仮説は深く考察すれば，かなり内容が豊富である．さらに一般的に (13.3.7) で示したように ω が $\beta_{k'+1}=\cdots=\beta_k=0$ の場合よりも，仮説 $\mathscr{H}'$; $\beta_{k'+1}=\beta_{k'+1,0},\cdots,\beta_k=\beta_{k0}$ の検定を行ないたい．この場合，$(dF_n)_\omega$ を求めるのに新しく確率変数 $y'_\xi = (y_\xi - \beta_{k'+1,0}x_{k'+1\xi} - \cdots - \beta_{k0}x_{k\xi})$ を導入しよう．この結果，正規分布 $N(\beta_1x_{1\xi}+\cdots+\beta_{k'}x_{k'\xi},\sigma^2)$ を持つ独立な確率変数 y'_ξ, $\xi=1,\cdots,n$ が得られ，仮説 $\mathscr{H}'$ は ω を $\beta_{k'+1}=\beta_{k'+1,0},\cdots,\beta_k=\beta_{k0}$ に変え，Ω はそのままにした (13.3.7) と同じである．尤度比 $\lambda_{\mathscr{H}'}$

は行列 $\|a_{i'_0j'_0}\|$, $\|a_{i'j'}\|$ のなかの y_ξ を y'_ξ で置き換えた $\lambda_{\mathscr{H}}$ とまったく構造が一致する．$\mathscr{H}'$ が真のときの $\lambda_{\mathscr{H}'}$ に関する分布論は $\mathscr{H}$ が真のときの $\lambda_{\mathscr{H}}$ に関するそれとまったく同じである．

今度は，$\beta_1,\cdots,\beta_k$ の間に次の形をしたある 1 次独立な線形制約があるという仮説 $\mathscr{H}''$ の検定をしたい．

$$\sum_{i=1}^k c_{iu}\beta_i = \gamma_{u0}, \qquad u=k'+1,\cdots,k.$$

ここに c_{iu} と γ_{u0} はある特定の数である．この場合，回帰関数 $\beta_1x_{1\xi}+\cdots+\beta_kx_{k\xi}$ に回帰係数の変換

$$\beta_1=\gamma_1,\cdots,\beta_{k'}=\gamma_{k'},$$
$$\sum_{i=1}^k c_{iu}\beta_i=\gamma_u, \qquad u=k'+1,\cdots,k$$

を行なうと $\gamma_1z_{1\xi}+\cdots+\gamma_kz_{k\xi}$ なる形をした回帰関数が得られる．ただし $z_{1\xi},\cdots,z_{k\xi}$ は c_{pu} に依存する係数を持つ $x_{1\eta},\cdots,x_{k\eta}$, $\eta=1,\cdots,n$ の線形関数である．よって仮説 $\mathscr{H}''$ は上述の仮説 $\mathscr{H}'$ の形に帰着する．すなわち $\beta_1,\cdots,\beta_k$ の代わりに $\gamma_1,\cdots,\gamma_k$ を，$x_{1\xi},\cdots,x_{k\xi}$ の代わりに $z_{1\xi},\cdots,z_{k\xi}$, $\xi=1,\cdots,n$ を用いているにすぎない．$\mathscr{H}''$ が真のときの $\lambda_{\mathscr{H}''}$ に関する標本論は $\mathscr{H}$ が真のときの $\lambda_{\mathscr{H}}$ に関するそれとまったく同じである．

13.4　大標本における尤度比の漸近分布

さて，大標本問題に話を戻そう．まず単純仮説

(13.4.1)
$$\mathscr{H}(\theta_0\,;\Omega)$$

の場合を考えよう．本節ではこれを $\mathscr{H}$ で表わす．ここで，θ は 1 次元のパラメータ，パラメータ空間 Ω は θ_0 を含んでいるある（開）区間 Ω_0 を含むような実数軸 R_1 上の区間とする．大標本論では尤度比検定の漸近的標本論において本質的な役割を果たすのは Ω のほんの一部の Ω_0 だけである．大きさ n の標本に対して，θ_0 に対する最尤推定量 $\hat\theta_n$ が存在して，推定量の列 $\hat\theta_n$, $n=1,2,\cdots$ は θ_0 に確率収束すると仮定する．したがって尤度比の対応する列は

(13.4.2)
$$\lambda_{\mathscr{H}}=\frac{\prod_{\xi=1}^{n} dF(x_\xi\,;\,\theta_0)}{\prod_{\xi=1}^{n} dF(x_\xi\,;\,\hat{\theta}_n)},\qquad n=1,2,\cdots$$

で与えられる．

以下に述べる条件の下では $\mathscr{H}$ が真のとき，

$$-2\log\lambda_{\mathscr{H}},\qquad n=1,2,\cdots$$

はカイ 2 乗分布 $C(1)$ に法則収束することを示そう．

$F(x\,;\,\theta)$ は Ω_0 において，θ の 1 次および 2 次導関数に関して正則であると仮定する [12.1 節を見よ]．この条件の下では，

(13.4.3)
$$H(\theta_0,\theta)=\int_{-\infty}^{\infty}\log dF(x\,;\,\theta)\,dF(x\,;\,\theta_0)$$

は積分記号の下で θ に関して 2 回微分可能である．よって

(13.4.4)
$$H'(\theta_0,\theta_0)=0,\qquad H''(\theta_0,\theta_0)=-B^2(\theta_0,\theta_0)$$

が得られる．ただし $B^2(\theta_0,\theta_0)$ は (12.1.8) に定義してある．

ゆえに $H(\theta_0,\theta)$ は $\theta=\theta_0$ で相対最大値 $H(\theta_0,\theta_0)$ を持つ．$H(\theta_0,\theta)$ は Kullback (1959) が展開した統計的情報理論では重要な役割を果たしている．一般に，$F_1(x), F_2(x), G(x)$ が 3 つの c.d.f. で，互いに絶対連続であれば，$G(x)$ からの観測あたりの $F_2(x)$ に対して $F_1(x)$ を識別するクルバックの**平均情報**，あるいは**情報積分**は $I(1\,;\,2)=\int_{-\infty}^{\infty}\log\dfrac{dF_1(x)}{dF_2(x)}\,dG(x)$ で定義される．よって

$$H(\theta_0,\theta_0)-H(\theta_0,\theta_1)=\int_{-\infty}^{\infty}\log\frac{dF(x\,;\,\theta_0)}{dF(x\,;\,\theta_1)}\,dF(x\,;\,\theta_0)$$

はクルバックの意味で，$F(x\,;\,\theta_0)$ からの観測あたりの $F(x\,;\,\theta_1)$ に対して $F(x\,;\,\theta_0)$ を識別する情報積分である．

注意　関数 $H(\theta,\theta)$ の特別な場合が統計力学に出てくる．x はその成分が与えられた気体分子の座標とモーメントを表わす点であるとしよう．x の空間は系の**相空間**といわれる．$F(x\,;\,t)$ が時間 t でのこの系の分子全体における x の c.d.f. を表わすならば，積分 $H(t,t)$ は (13.4.3) の $F(x\,;\,t)$ を用いて得られる．$H(t,t)$ は $-\log P$ (加法的定数をのぞいて) を近似するために Boltzmann (1910) により最初に導入された H 関数である．ただし

$$P=\frac{n!}{n_1!\,n_2!\cdots n_N!}\left(\frac{1}{N}\right)^n$$

は系の n 個の分子に対して，時刻 t において胞 $E_i,\ i=1,\cdots,N$ に n_i 個の分子が存在する確率である．ただし相空間における "胞" $E_1,\cdots,E_N$ は等確率 $\dfrac{1}{N}$ を持ち互いに素で，しかも全相空間を構成するようになっている．$H(t,t)$ は本質的には "もっとも起りそうな" あるいは "平衡な" 状態からの系の偏差の尺度である．

$H(\theta_0,\theta)$ 形の関数はまた Shannon (1948) により展開された通信理論でも重要な役割を果たしている．

12.2 (c) 節では，$B^2(\theta_0,\theta_0)$ はフィッシャーにより，$F(x\,;\,\theta_0)$ からの観測あたりの θ_0 に通ずる情報量として定義されていた．したがって，(13.4.4) に注意すれば，フィッシャーの観測あたりの情報量は，$\theta=\theta_0$ におけるクルバックの情報積分 $[H(\theta_0,\theta_0)-H(\theta_0,\theta)]$ の 2 次導関数であることがわかる．他方，$B^2(\theta_0,\theta_0)$ は本質的には $\theta=\theta_0$ で最大値をとる $H(\theta_0,\theta)$ の曲率の尺度ともなっている．まとめると，

13.4.1　$F(x\,;\,\theta)$ が Ω_0 において θ の 2 次導関数に関して正則であれば，$H(\theta_0,\theta)$ は $\theta=\theta_0$ で相対最大値 $H(\theta_0,\theta_0)$ を持つ．$\theta=\theta_0$ における $[H(\theta_0,\theta_0)-H(\theta_0,\theta)]$ (クルバックの情報積分) の 2 次導関数は $B^2(\theta_0,\theta_0)$ である．$B^2(\theta_0,\theta_0)$ は $F(x\,;\,\theta_0)$ からの観測あたりの θ_0 に通ずるフィッシャーの情報量である．

c.d.f. $F(x,\theta_0)$ からの標本に対して

(13.4.5)
$$\frac{1}{n}\sum_{\xi=1}^{n}\log dF(x_\xi\,;\,\hat{\theta})$$

は平均 $H(\theta_0,\theta_0)$ を持つ分布からの大きさ n の標本平均である．そして **9.1.1** より，この標本平均は $H(\theta_0,\theta_0)$ に確率収束する．ただし，$H(\theta_0,\theta_0)$ が存在するのは $F(x\,;\,\theta)$ が Ω_0 において θ の 1 次導関数に関して正則なときである．また，この条件の下では，**12.3.2** で見たように，$n\longrightarrow\infty$ のとき，$\hat{\theta}(x_1,\cdots,x_n),\ n=1,2,\cdots$ は θ_0 に確率収束する．さらに，**4.3.8** から

(13.4.6)
$$\frac{1}{n}\sum_{\xi=1}^{n}\log F(x_\xi\,;\,\hat{\theta})$$

は $n\longrightarrow\infty$ のとき $H(\theta_0,\theta_0)$ に確率収束する．定理にまとめると

13.4.2　$(x_1,\cdots,x_n)$ を $F(x\,;\,\theta_0)$ からの標本とする．ただし $F(x\,;\,\theta)$ は Ω_0 にお

いて θ の 1 次導関数に関して正則である．このとき

$$\frac{1}{n}\sum_{\xi=1}^{n}\log dF(x_\xi\,;\theta_0),\qquad \frac{1}{n}\sum_{\xi=1}^{n}F(x_\xi\,;\hat\theta)$$

は $n\longrightarrow\infty$ のとき，ともに $H(\theta_0,\theta_0)$ に確率収束する．

$\hat\theta$ が漸近的に正規分布しているという強い条件を満足しているときは，$\mathscr{H}$ が真のとき $-2\log\lambda_{\mathscr{H}}$ の漸近的分布に関する次の結果を得る．

13.4.3　$\hat\theta$ が **12.3.3** に従って漸近的に正規分布すれば，

(13.4.7)
$$\lim_{n\to\infty}P(-2\log\lambda_{\mathscr{H}}<\chi^2|\theta_0)=\frac{1}{\sqrt{2\pi}}\int_0^{\chi^2}u^{-\frac{1}{2}}e^{-\frac{1}{2}u}\,du.$$

すなわち，$\mathscr{H}$ が真のとき，$-2\log\lambda_{\mathscr{H}}$，$n=1,2,\cdots$ はカイ 2 乗分布 $C(1)$ に法則収束する．

13.4.3 の証明には，まず **12.3.3** の仮定を参照する．$\hat\theta$ は $n\longrightarrow\infty$ のとき θ_0 に概収束するので，任意の $\varepsilon>0$ に対して次の不等式が成り立つ確率が $1-\varepsilon$ を越えるような n_ε が存在する．$n>n_\varepsilon$ なるすべての n に対して

(13.4.8)
$$\sum_{\xi=1}^{n}\log dF(x_\xi\,;\theta_0)=\sum_{\xi=1}^{n}\log dF(x_\xi\,;\hat\theta)+\frac{1}{2}\sum_{\xi=1}^{n}\left[\frac{\partial^2}{\partial\theta^2}\log dF(x_\xi\,;\theta)\right]_{\theta=\theta^*}(\theta_0-\hat\theta)^2$$

が成り立つ．ただし，ここで θ^* は $|\theta_0-\theta^*|<|\theta_0-\hat\theta|$ なる確率変数である．さて，(13.4.8) は

(13.4.9)
$$-2\sum_{\xi=1}^{n}\log\left(\frac{dF(x_\xi\,;\theta_0)}{dF(x_\xi\,;\hat\theta)}\right)=-\frac{1}{n}\sum_{\xi=1}^{n}\left[\frac{\partial^2}{\partial\theta^2}\log dF(x_\xi\,;\theta)\right]_{\theta=\theta^*}[\sqrt{n}(\theta_0-\hat\theta)]^2$$

に書き直せる．(13.4.9) の左辺は $-2\log\lambda_{\mathscr{H}}$ になる．$n>n_\varepsilon$ なるすべての n に対して (13.4.9) が成り立つ確率が $1-\varepsilon$ を越えているから，(13.4.9) の左辺および右辺はともにカイ 2 乗分布 $C(1)$ に法則収束する確率変数列になる．ただし，$n=1,2,\cdots$．

(13.4.10)
$$-\frac{1}{n}\sum_{\xi=1}^{n}\left[\frac{\partial^2}{\partial\theta^2}\log dF(x_\xi\,;\theta)\right]_{\theta=\theta^*}$$

は $B^2(\theta_0,\theta_0)$ に確率収束する．さらに，$\sqrt{n}B(\theta_0,\theta_0)(\theta_0-\hat\theta)$ は $N(0,1)$ に法則収束する．

ゆえに，(13.4.9) の右辺の確率変数列はカイ 2 乗分布 $C(1)$ に法則収束する．よって (13.4.9) の左辺，すなわち，$-2\log\lambda_{\mathscr{H}}$ はカイ 2 乗分布 $C(1)$ を持つ確率変数に確率収束する．これで **13.4.3** の証明は終わる．

13.5　尤度比検定の一致性

W_α は

(13.5.1)
$$-2\log\lambda_{\mathscr{H}}>\chi^2_\alpha$$

を満たす R_n における集合であるとしよう．ただし，$\mathscr{H}$ は仮説 $\mathscr{H}(\theta_0\,;\Omega_0)$ で，χ^2_α は $P(\chi^2>\chi^2_\alpha)=\alpha$ なる数である．すなわち，この確率はカイ 2 乗分布 $C(1)$ から算出される．$F(x\,;\theta)$ が Ω_0 においてその θ の 2 次導関数に関して正則であれば，**13.4.3** から次式が成り立つ．

(13.5.2)
$$\lim_{n\to\infty}P(W_\alpha|\theta_0)=\alpha.$$

さて，$\theta_1\neq\theta_0$ を Ω_0 の任意の点として，

(13.5.3)
$$\lim_{n\to\infty}P(W_\alpha|\theta_1)$$

を考えよう．このとき

(13.5.4)
$$\lambda_{\mathscr{H}}=\lambda'_{\mathscr{H}}\cdot\nu$$

と書ける．ただし

$$\lambda'_{\mathscr{H}}=\prod_{\xi=1}^{n}\left(\frac{dF(x_\xi\,;\theta_1)}{dF(x_\xi\,;\hat\theta)}\right),\qquad \nu=\prod_{\xi=1}^{n}\left(\frac{dF(x_\xi\,;\theta_0)}{dF(x_\xi\,;\theta_1)}\right).$$

(13.5.5)
$$-2\log\lambda'_{\mathscr{H}}=u_n,\qquad -2\log\nu=nv_n$$

とすれば，W_α は

(13.5.6)
$$u_n+nv_n>\chi^2_\alpha$$

なる R_n における集合である．

θ の真値が θ_1 であれば，**13.4.3** の θ_0 を θ_1 で置き換えると，u_n，$n=1,2,\cdots$ はカイ 2 乗分布 $C(1)$ に法則収束する確率変数列であることがわかる．そして

(13.5.7) $$v_n = 2\Big[\frac{1}{n}\sum_{\xi=1}^{n}\log dF(x_\xi\,;\theta_1) - \frac{1}{n}\sum_{\xi=1}^{n}\log dF(x_\xi\,;\theta_0)\Big]$$

と書ける.

9.1.1 から, θ_1 が θ の真値であれば, $\frac{1}{n}\sum_{\xi=1}^{n}\log dF(x_\xi\,;\theta_1)$, $\frac{1}{n}\sum_{\xi=1}^{n}\log dF(x_\xi\,;\theta_0)$ はそれぞれ $H(\theta_1,\theta_1)$, $H(\theta_1,\theta_0)$ に確率収束する. しかし, **13.4.1** より, $H(\theta_1,\theta_1) > H(\theta_1,\theta_0)$. それゆえ確率変数 v_n, $n=1,2,\cdots$ は正の数 $v_0 = 2[H(\theta_1,\theta_1) - H(\theta_1,\theta_0)]$ に確率収束する. 一方, u_n, $n=1,2,\cdots$ はカイ 2 乗分布 $C(1)$ に法則収束するので,

(13.5.8) $$\lim_{n\to\infty} P(u_n + nv_n > \chi_\alpha^2) = 1$$

となる. これは

(13.5.9) $$\lim_{n\to\infty} P(W_\alpha | \theta = \theta_1) = 1$$

と同値である. すなわち, W_α は $\mathscr{H}(\theta_0\,;\Omega)$ に対する一致検定である.

それゆえ, Ω における θ に対して検出力関数列

(13.5.10) $$P(W_\alpha|\theta),\qquad n=1,2,\cdots$$

は関数

(13.5.11) $$\zeta(\theta) = \begin{cases}\alpha, & \theta = \theta_0\\ 1, & \theta \neq \theta_0\end{cases}$$

に収束する. これは, 十分大きな n に対して W_α (すなわち $-2\log\lambda_{\mathscr{H}}$) は $\mathscr{H}$ に対する一致検定であることを意味している. 要約すると次のようになる.

13.5.1 $F(x\,;\theta)$ が θ の 1 次, 2 次導関数に関してともに正則ならば, (13.5.1) で定義される尤度検定は仮説 $\mathscr{H}(\theta_0\,;\Omega)$ に対する一致検定である.

13.6 尤度比検定の漸近的検出力

13.4.3 から, 大標本に対する尤度比 $\lambda_{\mathscr{H}}$ は仮説 $\mathscr{H}(\theta_0\,;\Omega)$ の検定を定めることがわかる. この仮説に対する第 1 種の誤りの確率は自由度 1 のカイ 2 乗分布から計算される. また **13.5.1** より. この検定は一致検定である. すなわち, $n\longrightarrow\infty$ のとき, 検定の検出力関数は (13.5.11) で定義された $\zeta(\theta)$ に収束する. さて, ある条件の下では, 大きさ α の $\mathscr{H}(\theta_0\,;\Omega)$ に対する任意の一致検定の検出力関数は, $\lambda_{\mathscr{H}}$ から定められる検定より速く $\zeta(\theta)$ に収束しないことを示そう.

この目的のためには, 12.5 節で定義された正則な推定関数 $g_n(x_1,\cdots,x_n\,;\theta)$ に戻って考えれば十分である. すなわち 12.5 節では, $\mathscr{H}$ が真ならば, (12.5.11) と (12.5.12) で与えられる 2 つの確率変数列は $n\longrightarrow\infty$ のとき, ともに正規分布 $N(0,1)$ に法則収束する性質を持つことを見た. よって列

(13.6.1) $$ng_n^2(x_1,\cdots,x_n\,;\theta_0),\qquad n=1,2,\cdots$$

はカイ 2 乗分布 $C(1)$ に法則収束する. W_α^* を

(13.6.2) $$ng_n^2(x_1,\cdots,x_n\,;\theta_0) > \chi_\alpha^2$$

で定義される検定としよう. ここに, χ_α^2 は (13.5.1) で定義したのと同じ数である. ゆえに

(13.6.3) $$\lim_{n\to\infty} P(W_\alpha^*|\theta_0) = \alpha$$

が成り立つ. さて, $g_n(x_1,\cdots,x_n\,;\theta)$ が θ_0 の正則な推定関数という性質から, θ の真値が $\theta_1 \neq \theta_0$ のとき (ただし θ_1 は Ω_0 の点),

(13.6.4) $$ng_n^2(x_1,\cdots,x_n\,;\theta_0),\qquad n=1,2,\cdots$$

および

(13.6.5) $$w_n^* = \{g_n'(x_1,\cdots,x_n\,;\tilde{\theta}^*)[\sqrt{n}(\theta_1-\tilde{\theta}) + \sqrt{n}(\theta_0-\theta_1)]\}^2,\quad n=1,2,\cdots$$

は一方が法則収束すれば他方も法則収束する. しかし

(13.6.6) $$w_n^* = u_n^* + nv_n^*$$

と書ける. ここに

(13.6.7) $$u_n^* = [g_n'(x_1,\cdots,x_n\,;\tilde{\theta}^*)\sqrt{n}(\theta_1-\tilde{\theta})]^2$$
$$v_n^* = \frac{1}{n}[w_n^* - u_n^*].$$

θ の真値は θ_1 だから, 列 u_n^*, $n=1,2,\cdots$ はカイ 2 乗分布 $C(1)$ に法則収束し, v_n^*, $n=1,2,\cdots$ は定数 v_0^* に確率収束する. ただし

(13.6.8) $$v_0^* = B^{*2}(\theta_1,\theta_1)(\theta_0-\theta_1)^2.$$

さて, $\theta_1 \neq \theta_0$ かつ $B^{*2}(\theta_1,\theta_1) > 0$ であるので $v_0^* > 0$, よって

(13.6.9) $$\lim_{n\to\infty} P(u_n^* + nv_n^* > \chi_\alpha^2) = 1.$$

これは

(13.6.10) $$\lim_{n\to\infty} P(W_\alpha^*|\theta = \theta_1) = 1$$

と同値である．しかし，(13.6.3)，(13.6.10) の両式は不等式 (13.6.2) で与えられた検定が $\mathscr{H}$ の一致検定であることを示している．

不等式 (13.5.1) で定義された尤度比検定 W_α の検出力関数と不等式 (13.6.2) で定義された任意の検定 W_α^* を比較する問題は，両方とも一致検定であり，したがって，$n \longrightarrow \infty$ のとき，両者は (13.5.11) で定義された同一の極限検出力関数 $\zeta(\theta)$ を持つ，ということから複雑になる．これを明らかにするには次に新しい概念を導入して用いよう．

もし W_α が標本 $(x_1, \cdots, x_n)$ に基づく $\mathscr{H}$ の検定で，$n \longrightarrow \infty$ のときの標本の極限検出力関数 $\eta(\theta)$ は Ω_0 に含まれる θ に対して次式のように定義されているとする．

(13.6.11)
$$\eta(\theta_0) = \alpha, \qquad \alpha < \eta(\theta) < 1, \quad \theta \neq \theta_0.$$

このとき，W_α を有意水準 α における $\mathscr{H}(\theta_0 ; \Omega)$ に対する**漸近的不偏検定**という．検定の漸近的不偏性は一致性よりも検出力に関して弱い条件であることを注意しておこう．

$W_{1\alpha}$，$W_{2\alpha}$ を $\mathscr{H}$ に対する大きさ α の 2 つの漸近的不偏検定で，その Ω_0 における θ に対する漸近的検出力関数をそれぞれ $\eta_1(\theta)$，$\eta_2(\theta)$ とする．

もし $\eta_1(\theta_0) = \eta_2(\theta_0) = \alpha$ で

(13.6.12)
$$\eta_1(\theta) \geqslant \eta_2(\theta)$$

ならば，$W_{1\alpha}$ は $\mathscr{H}$ の検定に対して $W_{2\alpha}$ より**漸近的により強力**[*] であるという．もし (13.6.12) の等式が Ω_0 のいたるところで成り立つならば，$W_{1\alpha}$ と $W_{2\alpha}$ は仮説 $\mathscr{H}$ に対する大きさ α の**同値な漸近的不偏検定**であるという．

さて問題をそれぞれ (13.5.1)，(13.6.2) で定義された W_α，W_α^* の検出力の比較に戻そう．それぞれ不等式

(13.6.13)
$$u_n + cv_n > \chi_\alpha^2, \qquad u_n^* + cv_n^* > \chi_\alpha^2$$

で定義される検定，$W_{c\alpha}$，$W_{c\alpha}^*$ を考えよう．ここで $c > 0$ は任意である．u_n，u_n^* はともに非負だから，$n > c$ に対して $W_{c\alpha} \subset W_\alpha$，$W_{c\alpha}^* \subset W_\alpha^*$ は明らかである．

すべての $c > 0$ に対して，$W_{c\alpha}$ は仮説 $\mathscr{H}$ の検定に対して，W_α^* より漸近的に強力であることを示そう．この場合に，Wald (1941 *a*) は，W_α は仮説 $\mathscr{H}$ の検定に対しては W_α^* よりも漸近的により有力[**] であると述べている．

もし $W_{c\alpha}$ と $W_{c\alpha}^*$ がすべての $c > 0$ に対して漸近的に同値ならば，W_α と W_α^* は仮説 $\mathscr{H}$ の検定に対しては，**漸近的に等強力**であるという．

[*] 厳密に述べれば，(13.6.12) で $>$ だけが成り立つ θ の値が少なくとも 1 つある場合以外は，漸近的に少なくとも同程度に強力という用語を用いるべきであるが，誤解がなければ，この用語を用いる．

[**] 原書では *asymptotically more stringent* である．[訳注]

漸近的に等強力な検定に関する重要な結果は次のように表わされる．

13.6.1　W_α を (13.5.1) で定義され尤度比検定とし，W_α^* を，(13.6.2) の $g_n(x_1, \cdots, x_n ; \theta_0)$ を $h_n(x_1, \cdots, x_n ; \theta_0)$ で置き換えて定義されたものとする．ただし $h_n(x_1, \cdots, x_n ; \theta_0)$ は (12.5.1) で定義されている．このとき，c.d.f. $F(x ; \theta)$ が Ω_0 において θ の 2 次導関数に関して正則であれば，W_α と W_α^* は仮説 $\mathscr{H}$ の検定に対して漸近的に等強力である．

13.5 節で，(13.6.13) における (u_n, v_n) は，分布が $n \longrightarrow \infty$ のとき，(u, v) 平面の退化した分布に収束する確率変数の組であることを示した．この分布は半直線 $v = v_0$ $(u > 0)$ に沿ったカイ 2 乗分布 $C(1)$ であり，

(13.6.14)
$$v_0 = 2[H(\theta_1, \theta_1) - H(\theta_1, \theta_0)]$$

である．同様に (u_n^*, v_n^*) も分布が (u, v) 平面の退化した分布に収束する確率変数の組である．この分布は半直線 $v = v_0^*$ $(u > 0)$ に沿ったカイ 2 乗分布 $C(1)$ である．ただし

(13.6.15)
$$v_0^* = B^{*2}(\theta_1, \theta_1)(\theta_0 - \theta_1)^2.$$

しかし

(13.6.16)
$$H(\theta_1, \theta_0) = H(\theta_1, \theta_1) + H'(\theta_1, \theta_1)(\theta_0 - \theta_1) + \frac{1}{2}H''(\theta_1, \theta_1^*)(\theta_0 - \theta_1)^2.$$

ここに $|\theta_0 - \theta_1^*| < |\theta_0 - \theta_1|$．$H'(\theta_1, \theta_1) = 0$ より

(13.6.17)
$$2[H(\theta_1, \theta_1) - H(\theta_1, \theta_0)] = -H''(\theta_1, \theta_1^*)(\theta_0 - \theta_1)^2$$

しかし

$$-H''(\theta_1, \theta_1) = B^2(\theta_1, \theta_1).$$

また $F(x ; \theta)$ は Ω_0 において θ の 2 次導関数に関して正則だから，$\theta = \theta_1$ では

(13.6.18)
$$\frac{B^{*2}(\theta_1, \theta_1)}{B^2(\theta_1, \theta_1)} \leqslant 1$$

という結論を得る．もちろんこれは (12.5.21) で $\theta = \theta_0$ としたのと同様である．それゆえ $H''(\theta_1, \theta)$ は Ω_0 上で θ の連続関数であるので，たがいに十分に近い θ_0 と θ_1 に対して，(13.6.18) は $B^2(\theta_1, \theta_1)$ を $-H''(\theta_1, \theta_1^*)$ で置き換えても成り立つ．よって

(13.6.19)
$$0 < v_0^* \leqslant v_0.$$

さて，c の固定した値に対する $W_{c\alpha}$，$W_{c\alpha}^*$ の検出力関数

(13.6.20)
$$P(u_n + cv_n > \chi_\alpha^2 | \theta_1), \qquad P(u_n^* + cv_n^* > \chi_\alpha^2 | \theta_1)$$

を θ_1 の関数として考えよう．$n \longrightarrow \infty$ のとき，これらの関数の極限，すなわち $W_{c\alpha}$，$W^*_{c\alpha}$ の漸近的検定力関数をそれぞれ

(13.6.21)
$$\eta_c(\theta_1), \qquad \eta^*_c(\theta_1)$$

で表わす．

u_n，u^*_n は非負であり，(u_n, v_n) と (u^*_n, v^*_n) の分布は $n \longrightarrow \infty$ のとき (u, v) 平面の退化した分布に収束する．ただしこの退化分布はそれぞれ半直線，$v = v_0$，$u > 0$，$v = v^*_0$，$u > 0$ に沿ったカイ 2 乗分布 $C(1)$ である．したがって，

(13.6.22)

（i）$\eta_c(\theta_1) = \eta^*_c(\theta_1) = 1$，　$v^*_0 \geqq \dfrac{\chi^2_\alpha}{c}$ なる θ_1 の値に対して

（ii）$1 > \eta_c(\theta_1) > \eta^*_c(\theta_1) > \alpha$，　$0 < v^*_0 < \dfrac{\chi^2_\alpha}{c}$ なる θ_1 の値に対して

（iii）$\eta_c(\theta_0) = \eta^*_c(\theta_0) = \alpha$，　$\theta_1 = \theta_0$，すなわち $v^*_0 = 0$ なる θ_1 に対して

しかし，(13.6.8) から (13.6.22) の（i），(ii)，(iii) で示される θ_1 の値の 3 つの集合はそれぞれ，（i）θ_0 を含むある区間 Ω'_0 の外をのぞく，Ω_0 における θ_1 の値，（ii）Ω'_0 における $\theta_1 \neq \theta_0$ なる値，(iii) $\theta_1 = \theta_0$，からなっていることがわかる．

ゆえに，すべての $c > 0$ に対して漸近的検出力関数 $\eta_c(\theta)$，$\eta^*_c(\theta)$ は (13.6.11) を満足し，この対は (13.6.12) を満足するので，すべての $c > 0$ に対して仮説 $\mathscr{H}$ の検定には，$W^*_{c\alpha}$ が $W_{c\alpha}$ より漸近的に強力である．したがって，W_α は $\mathscr{H}$ の検定に対しては W^*_α より漸近的に強力である．

$v^*_0 = v_0$ ならば，$0 < v^*_0 < \chi^2_\alpha/c$ なる θ_1 の値に対して (13.6.22) の (ii) は
$$1 > \eta_c(\theta_1) = \eta^*_c(\theta_1) > \alpha$$
と置き換えた型が成り立つ．よって $W_{c\alpha}$ と $W^*_{c\alpha}$ はすべての $c > 0$ に対して同値な漸近的不偏検定になる．すなわち，W_α と W^*_α は仮説 $\mathscr{H}$ の検定に関して漸近的に等強力である．しかし $v^*_0 = v_0$ であるための必要十分条件は (13.6.18) の等号が成り立つことである．すなわち，W^*_α の定義のとき用いた正則推定関数 $g_n(x_1, \cdots, x_n ; \theta)$ を $h_n(x_1, \cdots, x_n ; \theta)$ で置き換えることである．しかし **13.6.1** から (13.6.2) で $h_n(x_1, \cdots, x_n ; \theta)$ を用いれば，検定 W^*_α は (13.5.1) で定義された検定 W_α と漸近的に等強力であることがわかる．

以上をまとめると次の重要な定理となる．

13.6.2　$(x_1, \cdots, x_n)$ を c.d.f. $F(x ; \theta)$ からの標本とする．ただし，$F(x ; \theta)$ は Ω_0 においてその θ の 2 次導関数に関して正則である．W_α を (13.5.1) で定義された大きさ α の尤度比検定とする．$g_n(x_1, \cdots, x_n ; \theta)$ は 12.5 節で定義され θ に対する正則な推定関数で，W^*_α を (13.6.2) で定義したように $g_n(x_1, \cdots, x_n ; \theta)$ に基づく大きさ α の検定とする．このとき，仮説 $\mathscr{H}(\theta_0 ; \Omega)$ の検定に対して W_α は W^*_α より漸近的に強力である．しかし，(13.6.2) の $g_n(x_1, \cdots, x_n ; \theta)$ を (12.5.1) で定義した $h_n(x_1, \cdots, x_n ; \theta)$ で置き換えれば，仮説 $\mathscr{H}(\theta_0 ; \Omega)$ の検定に関して W_α と W^*_α は漸近的に等強力になる．

13.7　単純仮説の尤度比検定

13.4，13.5，13.6 節では主に単純仮説に対する尤度比検定の大標本における漸近的性質について述べた．そこで取り扱った単純仮説ではパラメータ θ は 1 次元であった．事実，パラメータ空間 Ω の重要な部分は θ_0 を含む Ω における開区間 Ω_0 であり，この開区間を中心に理論を展開してきた．これまでに展開した漸近的性質は θ が r 次元パラメータ $(\theta_1, \cdots, \theta_r)$ の場合にも容易に拡張できる．それゆえ，13.4 から 13.6 節にわたる結果の r 次元への拡張は詳細には述べる必要はなかろう．今後，c.d.f. $F(x ; \theta)$ を持つ確率変数 x を 1 次元確率変数として取り扱おう．しかし，以下の結果は記号を変えるだけで k 次元確率変数に対してもそのまま成り立つ．

ここで考える仮説は，単純仮説
$$\mathscr{H}(\theta_0 ; \Omega_r)$$
であり，$\mathscr{H}_r$ で表わす．ここで θ_0 は点 $(\theta_{10}, \cdots, \theta_{r0})$ で，Ω_r は θ_0 を含む r 次元開区間 Ω_{r0} を含むようなユークリッド空間 R_r における任意の集合である．$\mathscr{H}_r$ に対する尤度比検定の漸近理論を扱う上で重要なのは，Ω_r の一部 Ω_{r0} である．

まず，**13.4.1** の r 次元パラメータ θ への拡張を考えよう．

13.7.1　θ が r 次元で，$F(x ; \theta)$ が Ω_{r0} においてその θ の 2 次偏導関数すべてに関して正則ならば，(12.1.17) で定義された $H(\theta_0, \theta)$ は $\theta = \theta_0$ で最大値 $H(\theta_0, \theta_0)$ を持つ．さらに，$\theta = \theta_0$ における $H(\theta_0, \theta)$ の 2 次偏導関数の行

列の負*)は (12.1.19) で定義される $\|B_{pq}(\theta_0,\theta_0)\|$, $p,q=1,\cdots,r$ になる．

13.7.1 の証明は，**13.4.1** を r 次元パラメータの場合に拡張したのと同じなので割愛する．r 次元の場合，$H(\theta_0,\theta_0)-H(\theta_0,\theta_1)$ は $F(x;\theta_0)$ からの観測あたりの，$F(x;\theta_1)$ に対して $F(x;\theta_0)$ を識別するクルバックの情報積分である．フィッシャーの，$F(x;\theta_0)$ からの観測あたりの θ_0 に通ずる情報マトリックス $\|B_{pq}(\theta_0,\theta_0)\|$ はクルバックの情報積分の $\theta_1=\theta_0$ における 2 次偏導関数の行列である．[Kullback と Leibler (1951) を参照せよ]．

13.4.2 の r 次元への拡張は，θ, θ_0, $\hat\theta$ を r 次元と解釈すれば改めていい換える必要はない．**13.4.3** の r 次元への拡張は次のようになる．

13.7.2　θ が r 次元で，$F(x;\theta)$ が θ の 2 次偏導関数すべてに関して正則で，仮説 $\mathscr{H}_r$ が真ならば

(13.7.1)
$$\lim_{n\to\infty}P(-2\log\lambda_{\mathscr{H}_r}<\chi^2)=\frac{1}{2^{\frac{1}{2}r}\Gamma\left(\frac{1}{2}r\right)}\int_0^{\chi^2}u^{\frac{1}{2}r-1}e^{-\frac{1}{2}u}\,du.$$

すなわち $-2\log\lambda_{\mathscr{H}_r}$ はカイ 2 乗分布 $C(r)$ に法則収束する．

この証明は **13.4.2** と同様なので読者に残しておこう．

定理 **13.5.1** は θ が r 次元の場合にも成り立つ．これは **13.5.1** の単なる拡張に過ぎないので割愛しよう．

13.6.1 を θ が r 次元の場合に拡張すれば，W_α と W_α^* は R_n において不等式

(13.7.2)
$$-2\log\lambda_{\mathscr{H}}>\chi_\alpha^2$$

(13.7.3)
$$U_n(x_1,\cdots,x_n;\theta_0)>\chi_\alpha^2$$

によって定義される検定になる．ここに $U_n(x_1,\cdots,x_n;\theta_0)$ は (12.9.1) で定義されており，χ_α^2 はカイ 2 乗分布 $C(r)$ の $100\alpha\%$ 点である．この証明は **13.6.1** と同じであるので読者にゆだねる．

定理 **13.6.2** の r 次元パラメータへの拡張も容易である．読者に演習問題に残しておこう．**13.6.2** で用いた W_α^* の r 次元への拡張は $V_n>\chi_\alpha^2$ なる標本空間の領域になる．ただし，V_n は (12.9.15) に定義されている．1 次元推定関数 $h_n(x_1,\cdots,x_n;\theta)$ の代わりには，もちろん (12.9.1) で定義したベクトル関数 $h_{pn}(x_1,\cdots,x_n;\theta)$，すなわち

(13.7.4)
$$h_{pn}(x_1,\cdots,x_n;\theta)=\frac{1}{n}S_{pn}(x_1,\cdots,x_n;\theta),\qquad p=1,\cdots,r$$

*) 行列 $X=\|x_{ij}\|$ に対して $-X=\|-x_{ij}\|$ を X の負という．[訳注]

を用いればよい．

13.8　複合仮説の尤度比検定

いくつかのパラメータを含む仮説を検定する場合，パラメータの部分空間 $\omega_{r'}$ が Ω_r の r' 次元ユークリッド空間の断面の場合がしばしば出てくる．すなわち $(\theta_1,\cdots,\theta_{r'},\theta_{r'+10},\cdots,\theta_{r0})$（ただし $\theta_{r'+10},\cdots,\theta_{r0}$ は固定されている）なる形をした点からなる筒集合の場合である．この複合仮説を $\mathscr{H}(\omega_{r'};\Omega_r)$ または単に $\mathscr{H}_{r-r'}$ で表わす．$\mathscr{H}_{r-r'}$ に対する尤度比検定 $\lambda_{\mathscr{H}_{r-r'}}$ は (13.3.5) による通常の方法で定義される．大標本において，Ω_r の重要な部分は真のパラメータ点 $(\theta_{10},\cdots,\theta_{r0})$ を含むある r 次元開区間 Ω_{r0} にはいる部分である．よって大標本に対しては $\mathscr{H}_{r-r'}$ は複合仮説 $\mathscr{H}(\omega_{r'0};\Omega_{r0})$ と考えられる．ただし $\omega_{r'0}=\omega_{r'}\cap\Omega_{r0}$．

尤度比検定に関する基本定理 [Wilks (1938 *a*)] は次のとおりである．

13.8.1　$(x_1,\cdots,x_n)$ を c.d.f. $F(x;\theta)$ からの標本とする．ここで θ は r 次元で，$F(x;\theta)$ は $\theta\in\Omega_{r0}$ に対してその θ の 2 次偏導関数のすべてに関して正則であるとする．このとき

(13.8.1)
$$\lim_{n\to\infty}P(-2\log\lambda_{\mathscr{H}_{r-r'}}<\chi^2|\theta\in\omega_{r'})=\frac{1}{2\Gamma\left(\frac{r-r'}{2}\right)}\int_0^{\chi^2}\left(\frac{u}{2}\right)^{[(r-r')/2]-1}e^{-\frac{1}{2}u}\,du$$

である．すなわち $\mathscr{H}_{r-r'}$ が真であれば，$-2\log\lambda_{\mathscr{H}_{r-r'}}$, $n=1,2,\cdots$ はカイ 2 乗分布 $C(r-r')$ を持つ確率変数に確率収束する．

証明はまず，$\hat\theta_1,\cdots,\hat\theta_r$ を **12.7.2** で定義したものとしよう．$\hat\theta_1',\cdots,\hat\theta_{r'}'$ を $\theta_1,\cdots,\theta_{r'}$ に関する

(13.8.2)
$$S_{p'n}(x_1,\cdots,x_n;\theta)=0,\qquad p'=1,\cdots,r'$$

の解とする．すなわち，$\hat\theta_1',\cdots,\hat\theta_{r'}'$ は，他の成分を $\theta_{r'+10},\cdots,\theta_{r0}$ に固定したときの $\theta_1,\cdots,\theta_{r'}$ に対する最尤推定量である．**12.7.3** の仮定の下では，$\hat\theta_1',\cdots,\hat\theta_{r'}'$ は n が大きいとき

(13.8.3)
$$N(\{\theta_{p'0}\};\|nB_{p'q'}\|^{-1}),\qquad p',q'=1,\cdots,r'$$

に従って漸近的に分布する．

さて (12.6.1) を参照して次の記号を用いよう．$\theta_0\in\omega_{r'}$ に対して

$$\frac{1}{\sqrt{n}}S_{pn}(x_1,\cdots,x_n\,;\theta_0)=\zeta_{np}$$

(13.8.4)
$$\sqrt{n}(\theta_{p0}-\hat{\theta}_p)=\eta_{np},\qquad \sqrt{n}(\theta_{p'0}-\hat{\theta}_{p'})=\eta'_{np'}$$
$$\frac{1}{n}S_{pqn}(x_1,\cdots,x_n,\theta)=A^{(n)}_{pq}(\theta)$$
$$p,q=1,\cdots,r,\qquad p',q'=1,\cdots,r'.$$

$F(x\,;\theta)$ の θ の 2 次偏導関数すべてに関する正則性により，任意の $\varepsilon>0$ に対して，n_ε が存在して $n>n_\varepsilon$ なるすべての n について次の 4 つの式がすべて成り立つ確率が $1-\varepsilon$ を越える．

(13.8.5)
$$\zeta_{np}=\sum_{q=1}^{r}A^{(n)}_{pq}(\theta^*)\eta_{nq},\qquad p=1,\cdots,r$$

(13.8.6)
$$\zeta_{np'}=\sum_{q'=1}^{r'}A'^{(n)}_{p'q'}(\theta^{*\prime})\eta'_{nq'},\qquad p'=1,\cdots,r'$$

(13.8.7)
$$\sum_{\xi=1}^{n}\log dF(x_\xi,\theta_0)=\sum_{\xi=1}^{n}\log dF(x_\xi,\hat{\theta}_1,\cdots,\hat{\theta}_r)+\frac{1}{2}\sum_{p,q=1}^{r}A^{(n)}_{pq}(\theta_1^*)\eta_{np}\eta_{nq}$$

(13.8.8)
$$\sum_{\xi=1}^{n}\log dF(x_\xi,\theta_0)=\sum_{\xi=1}^{n}\log dF(x_\xi,\hat{\theta}_1,\cdots,\hat{\theta}'_{r'},\theta_{r'+10},\cdots,\theta_{r0})+\frac{1}{2}\sum_{p',q'=1}^{r'}A^{(n)}_{p'q'}(\theta_1^{*\prime})\eta'_{np'}\eta'_{nq'}$$

ただし，θ^*, θ_1^* は θ_0 と $\hat{\theta}$ の間の Ω_{r0} における点で，$\theta^{*\prime}$, $\theta_1^{*\prime}$ は θ_0 と $(\hat{\theta}'_1,\cdots,\hat{\theta}'_{r'},\theta_{r'+10},\cdots,\theta_{r0})$ の間の $\omega_{r'}\cap\Omega_{r0}$ における点である．

η_{nq} と $\eta'_{nq'}$ は

(13.8.9)
$$\eta_{nq}=\sum_{p=1}^{r}A^{pq}_{(n)}(\theta^*)\zeta_{np},\qquad \eta'_{nq'}=\sum_{p'=1}^{r'}A'^{p'q'}_{(n)}(\theta^{*\prime})\zeta_{np'}$$

と表わせる．ここで

$$\|A^{pq}_{(n)}(\theta^*)\|=\|A^{(n)}_{pq}(\theta^*)\|^{-1},\quad \text{かつ}\quad \|A'^{p'q'}_{(n)}(\theta^{*\prime})\|=\|A'^{(n)}_{p'q'}(\theta^{*\prime})\|^{-1}.$$

(13.8.7) と (13.8.8) の左辺は同じで

(13.8.10)
$$-2\log\lambda_{\mathscr{H}_{r-r'}}=-2\left\{\sum_{\xi=1}^{n}\log dF(x_\xi,\hat{\theta}'_1,\cdots,\hat{\theta}'_{r'},\theta_{r'+10},\cdots,\theta_{r0})-\sum_{\xi=1}^{n}\log dF(x_\xi,\hat{\theta}_1,\cdots,\hat{\theta}_r)\right\}$$

である．よって，すべての $n>n_\varepsilon$ に対して，次の等式が成り立つ確率は $1-\varepsilon$ を越える．

(13.8.11)
$$-2\log\lambda_{\mathscr{H}_{r-r'}}=\sum_{p',q'=1}^{r'}A^{(n)}_{p'q'}(\theta_1^{*\prime})\eta'_{np'}\eta'_{nq'}-\sum_{p,q=1}^{r}A^{(n)}_{pq}(\theta_1^*)\eta_{np}\eta_{nq}.$$

ただし η_{np}, $\eta'_{np'}$ は等式 (13.8.9) により ζ_{np} で表わされる．

$\theta_0\in\omega_{r'}$ が θ の真値ならば，**12.7.1** から $(\zeta_{n1},\cdots,\zeta_{nr})$, $n=1,2,\cdots$ は $N(\{0\}\,;\|B_{pq}\|)$ に法則収束する確率変数列になる．さらに，行列 $\|A^{(n)}_{pq}(\theta^*)\|$, $\|A^{(n)}_{pq}(\theta_1^*)\|$ は $n\longrightarrow\infty$ のとき，ともに行列 $\|-B_{pq}\|$ に確率収束する．一方，行列 $\|A^{(n)}_{p'q'}(\theta^{*\prime})\|$, $\|A^{(n)}_{p'q'}(\theta_1^{*\prime})\|$ はともに $\|-B_{p'q'}\|$ に確率収束する．ただし B_{pq} は **12.7.1** で定義されている．

(13.8.11) の右辺を計算すると，(13.8.11) の両辺とも $n\longrightarrow\infty$ のとき次の確率変数 Q の分布に法則収束することがわかる．

(13.8.12)
$$Q=\sum_{p,q=1}^{r}B^{pq}y_py_q-\sum_{p',q'=1}^{r'}B^{p'q'}_{[r']}y_{p'}y_{q'}$$

ここで $\|B^{p'q'}_{[r']}\|=\|B_{p'q'}\|^{-1}$ で，$(y_1,\cdots,y_r)$ は分布 $N(\{0\},\|B_{pq}\|)$ を持つ．$\theta_0\in\omega_{r'}$ に対して，Q の分布はカイ 2 乗分布 $C(r-r')$ であることが示される．よって確率変数 $-2\log\lambda_{\mathscr{H}_{r-r'}}$ は Q と同じ分布に法則収束するので **13.8.1** の証明が終わる．

最後に，証明は省略するが，**13.8.1** の仮定の下では，$\mathscr{H}_{r-r'}$ に対する尤度比検定は一致検定，すなわち

$$\lim_{n\to\infty}P(-2\log\lambda_{\mathscr{H}_{r-r'}}>\chi^2_\alpha\,|\,\theta\in\Omega_r-\omega_{r'})=1$$

であることがわかる．(13.8.5) から (13.8.8) までにおけるスコア関数 $S_{pn}(x_1,\cdots,x_n\,;\theta)$ を用いないで，12.9 節で導入した正則な推定関数 $ng_{pn}(x_1,\cdots,x_n\,;\theta)$ を用いれば，$\mathscr{H}_{r-r'}$ に対する一致検定のクラスを構成することができる．しかし，このクラスのどの検定も，漸近的強力の概念を複合仮説にまでそのままの方法で拡張できたとしても，尤度比検定より漸近的に強力ではない．

問　　題

13.1　標本 $(x_1,\cdots,x_n)$ を $\theta e^{-\theta x}$, $x>0$ なる形の p.d.f. を持つ分布から抽出し，$\Omega=(0,+\infty)$ とする．ネイマン＝ピヤソンの定理を用いて，仮説 $\mathscr{H}(\theta_0\,;\theta_0\cup\theta_1)$ の検定に対して，$(x_1,\cdots,x_n)$ の標本空間において $\theta_1>\theta_0$ および $P(W_\alpha|\theta_0)=\alpha$ を満たす棄却域 W_α を定めよ．

この検定の検出力関数すなわち θ の関数 $P(W_\alpha|\theta)$ を求めよ．

13.2　n_1, $\bar{x}_1$, s_1^2 をそれぞれ分布 $N(\mu_1,\sigma^2)$ からの標本の大きさ，標本平均，標本分散とし，n_2, $\bar{x}_2$, s_2^2 を分布 $N(\mu_2,\sigma^2)$ からの独立な標本の同じ統計量とする．次で定義される複合仮説 $\mathscr{H}(\omega\,;\Omega)$ を考えよう．

Ω は $\sigma^2>0$ なる (μ_1,μ_2,σ^2) のユークリッド半空間 R_3.

ω は $\mu_1=\mu_2$ なる Ω の部分集合．

$\mathscr{H}$ の検定に対する尤度比はスチューデント比

$$t=\frac{(\bar{x}_1-\bar{x}_2)}{s\sqrt{1/n_1+1/n_2}}$$

ただし

$$s^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

と同値であることを示せ．ただし，$\mathscr{H}$ が真のとき，t は“スチューデント”分布 $S(n_1+n_2-2)$ を持つ．この検定の不偏性を示せ．

13.3　前述の問題は，n_i, $\bar{x}_i$, s_i^2, $i=1,\cdots,k$ がそれぞれ $N(\mu_i,\sigma^2)$, $i=1,\cdots,k$ からの k 個の独立な標本の大きさ，標本平均，標本分散の場合に一般化される．$\mathscr{H}(\omega\,;\Omega)$ を次のような統計的仮説とする．

Ω は $\sigma^2>0$ なる $(\mu_1,\cdots,\mu_k,\sigma^2)$ のユークリッド半空間 R_{k+1}.

ω は $\mu_1=\mu_2=\cdots=\mu_k$ なる Ω の部分集合．

このとき $\mathscr{H}(\omega\,;\Omega)$ に対する尤度比はスネディッカー比

$$\mathscr{F}=\frac{(n-k)\sum_{i=1}^{k}n_i(\bar{x}_i-\bar{x})^2}{k\sum_{i=1}^{k}(n_i-1)s_i^2}$$

に同値であることを示せ．ここに，$n=n_1+\cdots+n_k$, $\bar{x}=\dfrac{1}{n}\sum_{i=1}^{k}n_i\bar{x}_i$ で，また $\mathscr{H}(\omega\,;\Omega)$ が真のとき，$\mathscr{F}$ はスネディッカー分布 $S(k,n-k)$ を持つ．この検定が不偏であることを示せ．

13.4　s_1^2, s_2^2 をそれぞれ $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$ からの大きさ n_1, n_2 の標本分散，$\mathscr{H}(\omega\,;\Omega)$ を次のような統計的仮説とする．Ω は $\sigma_1^2>0$, $\sigma_2^2>0$ なる $(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2)$ のユークリッド $\dfrac{1}{4}$ 空間 R_4 で，ω は $\sigma_1^2=\sigma_2^2$ となる Ω の部分集合である．$\mathscr{H}(\omega\,;\Omega)$ に対する尤度比はスネディッカー比

$$\mathscr{F}=\frac{s_1^2}{s_2^2}$$

に同値であることを示せ．これは $\mathscr{H}(\omega\,;\Omega)$ が真のとき，スネディッカー分布 $S(n_1-1,n_2-1)$ を持つ．

13.5　x_1, x_2 は 2 項分布 $Bi(n_1,p_1)$, $Bi(n_2,p_2)$ を持つ独立な確率変数であるとする．$\mathscr{H}(\omega\,;\Omega)$ を次の統計的仮説とする．Ω は頂点 $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ を持つ単位矩形内のすべての可能な点 (p_1,p_2) からなる空間で，ω は $p_1=p_2$ なる Ω の部分集合である．$\mathscr{H}(\omega\,;\Omega)$ を検定する尤度比は

$$\lambda=\left(\frac{x}{n}\right)^{x}\left(1-\frac{x}{n}\right)^{n-x}\left(\frac{x_1}{n_1}\right)^{-x_1}\left(\frac{x_2}{n_2}\right)^{-x_2}\left(1-\frac{x_1}{n_1}\right)^{-n_1+x_1}\left(1-\frac{x_2}{n_2}\right)^{-n_2+x_2}$$

で与えられることを示せ．ただし $x=x_1+x_2$, $n=n_1+n_2$ で，$\mathscr{H}(\omega\,;\Omega)$ が真ならば，$n_1,n_2\longrightarrow\infty$ のときの $-2\log\lambda$ の極限分布はカイ 2 乗分布 $C(1)$ である．

13.6　前述の問題を $x_1,\cdots,x_k$ が 2 項分布 $Bi(n_1,p_1),\cdots,Bi(n_k,p_k)$ からの独立な確率変数である場合に一般化して，その解を求めよ．

13.7　**$r\times s$ 分割表における独立性の検定**　$(n_{ij},\ i=1,\cdots,r,\ j=1,\cdots,s)$ を多項分布

$$\frac{n!}{\prod_{j=1}^{s}\prod_{i=1}^{r}n_{ij}!}\prod_{j=1}^{s}\prod_{i=1}^{r}p_{ij}^{n_{ij}}$$

からの $(rs-1)$ 次元確率変数とする．ここに，$\sum_j\sum_i n_{ij}=n$（ただし各 $p_{ij}>0$ で $\sum_j\sum_i p_{ij}=1$）．$\mathscr{H}(\omega\,;\Omega)$ を次の統計的仮説とする．Ω は p_{ij} のすべての可能な値の空間で，ω は $p_{ij}=p_iq_j$（ただし $\sum_i p_i=\sum_j q_j=1$）なる Ω の部分集合である．$\mathscr{H}(\omega\,;\Omega)$ に対する尤度比は

$$\lambda=\frac{\prod_{i=1}^{r}\left(\frac{n_{i\cdot}}{n}\right)^{n_{i\cdot}}\prod_{j=1}^{s}\left(\frac{n_{\cdot j}}{n}\right)^{n_{\cdot j}}}{\prod_{j=1}^{s}\prod_{i=1}^{r}\left(\frac{n_{ij}}{n}\right)^{n_{ij}}}$$

で与えられることを示せ．ただし

$$n_{i\cdot}=\sum_j n_{ij},\qquad n_{\cdot j}=\sum_i n_{ij}$$

で，$\mathscr{H}(\omega\,;\Omega)$ が真のとき，$-2\log\lambda$ の極限分布はカイ 2 乗分布 $C((r-1)(s-1))$ になることを示せ．

13.8（続き）　g を次で定義する.

$$g=\sum_{j=1}^{s}\sum_{i=1}^{s}\left(n_{ij}-\frac{n_{i}.n._{j}}{n}\right)^2\Big/\left(\frac{n_{i}.n._{j}}{n}\right)$$

ただし

$$n_{i}.=\sum_{j}n_{ij},\qquad n._{j}=\sum_{i}n_{ij}$$

$\mathscr{H}$が真ならば，g と $-2\log\lambda$ はともに $n\longrightarrow\infty$ のとき，カイ 2 乗分布 $C((r-1)(s-1))$ に法則収束することを示せ.

13.9　$r\times s\times t$ 分割表における層の独立性　（n_{ijk}；$i=1,\cdots,r,\ j=1,\cdots,s,\ k=1,\cdots,t$）を

$$\frac{n\,!}{\prod_{k=1}^{t}\prod_{j=1}^{s}\prod_{i=1}^{r}n_{ijk}\,!}\prod_{k}\prod_{j}\prod_{i}p_{ijk}^{n_{ijk}}$$

の多項分布を持つ（$rst-1$）次元の確率変数とする.　ここで $\sum_{k}\sum_{j}\sum_{i}n_{ijk}=n$. $\sum_{k}\sum_{j}\sum_{i}p_{ijk}=1$ で $p_{ijk}>0$. $\mathscr{H}(\omega\,;\Omega)$ を，Ω が p_{ijk} のすべての可能な値の空間で，ω は $p_{ijk}=p_{ij}q_{k}$, $\sum_{j}\sum_{i}p_{ij}=\sum_{k}q_{k}=1$ なる Ω の部分集合であるような統計的仮説としよう. $\mathscr{H}$ の検定に対する尤度比 λ は

$$\lambda=\frac{\prod_{j}\prod_{i}\left(\frac{n_{ij.}}{n}\right)^{n_{ij.}}\prod_{k}\left(\frac{n_{..k}}{n}\right)^{n_{..k}}}{\prod_{k}\prod_{j}\prod_{i}\left(\frac{n_{ijk}}{n}\right)^{n_{ijk}}}$$

で与えられることを示せ. ここで $n_{ij.}=\sum_{k}n_{ijk}$, $n_{..k}=\sum_{j}\sum_{i}n_{ijk}$. $\mathscr{H}(\omega\,;\Omega)$ が真ならば，$n\longrightarrow\infty$ のとき，$-2\log\lambda$ の極限分布はカイ 2 乗分布 $C((rs-1)(t-1))$ になることを示せ.

13.10　$(x_{11},\cdots,x_{1n_1})$, $(x_{21},\cdots,x_{2n_2})$ はそれぞれポアソン分布 $Po(\mu_1)$, $Po(\mu_2)$ からの独立な標本であるとする. $\mathscr{H}(\omega\,;\Omega)$ を次の統計的仮説としよう. Ω は $\mu_1\mu_2$ 平面の第1象限の可能なすべての点 (μ_1,μ_2) の空間で，ω は Ω における $\mu_1=\mu_2$ なる直線である. $\mathscr{H}(\omega\,;\Omega)$ の尤度比 λ は

$$\lambda=\frac{\bar{x}^{n\bar{x}}}{\bar{x}_1^{n_1\bar{x}_1}\,\bar{x}_2^{n_2\bar{x}_2}}$$

で与えられることを示せ. ただし $\bar{x}_1$, $\bar{x}_2$ はそれぞれ 2 つの標本の平均で

$$\bar{x}=\frac{n_1\bar{x}_1+n_2\bar{x}_2}{n_1+n_2}$$

で $n=n_1+n_2$ である. $\mathscr{H}(\omega\,;\Omega)$ が真ならば，$n_1,n_2\longrightarrow\infty$ のとき $-2\log\lambda$ の極限分布はカイ 2 乗分布 $C(1)$ になることを示せ. さらに，k 標本の場合に一般化せよ.

13.11　$k-1$ 次元確率変数 $(n_1,\cdots,n_k)$ は多項分布

$$\frac{n\,!}{n_1\,!\cdots n_k\,!}\,p_1^{n_1}\cdots p_k^{n_k}$$

を持つ. ただし $n_1+\cdots+n_k=n$, $p_1+\cdots+p_k=1$ で $p_i>0$. 仮説 $\mathscr{H}(\omega\,;\Omega)$ の尤度比 λ は次式で与えられることを示せ. ここで Ω はすべての可能な p_i の空間で, ω は $p_1=p_{10},\cdots,p_k=p_{k0}$ なる Ω の部分集合である.

$$\lambda=\left[\left(\frac{p_{10}}{\hat{p}_1}\right)^{\hat{p}_1}\cdots\left(\frac{p_{k0}}{\hat{p}_k}\right)^{\hat{p}_k}\right]^n$$

ただし $\hat{p}_1=n_1/n,\cdots,\hat{p}_k=n_k/n$. $\mathscr{H}(\omega\,;\Omega)$ が真ならば, $n\longrightarrow\infty$ のとき，$-2\log\lambda$ の極限分布はカイ 2 乗分布 $C(k-1)$ になることを示せ.

13.12（続き）　$\mathscr{H}(\omega\,;\Omega)$ が真であれば, $n\longrightarrow\infty$ のとき $-2\log\lambda$, $\sum_{i=1}^{k}\frac{(n_i-np_{i0})^2}{np_{i0}}$ はともにカイ 2 乗分布 $C(r-1)$ に収束することを示せ.

13.13　$s_1^2,\cdots,s_k^2$ をそれぞれ，分布 $N(\mu_1,\sigma_1^2),\cdots,N(\mu_k,\ \sigma_k^2)$ からの大きさ $n_1,\cdots,n_k$ の独立な標本の分散とする. $\mathscr{H}(\omega\,;\Omega)$ を次の統計的仮説とする. Ω が $(\mu_1,\cdots,\mu_k,\sigma_1^2,\cdots,\sigma_k^2)$, $\sigma_1^2>0,\cdots,\sigma_k^2>0$ の可能な値すべてからなるユークリッド空間 R_{2k} の 2^{-k} 部分であり, ω が $\sigma_1^2=\cdots=\sigma_k^2$ なる Ω の部分集合である. このとき $\mathscr{H}(\omega\,;\Omega)$ の尤度比 λ は

$$\lambda=\left[\frac{(n_1-1)s_1^2}{n_1s_0^2}\right]^{\frac{1}{2}n_1}\cdots\left[\frac{(n_k-1)s_k^2}{n_ks_0^2}\right]^{\frac{1}{2}n_k}$$

で与えられることを示せ. ただし

$$s_0^2=\frac{(n_1-1)s_1^2+\cdots+(n_k-1)s_k^2}{n_1+\cdots+n_k}.$$

13.14　s^2 が $N(\mu,\sigma^2)$ からの大きさ n の標本分散であるとき, 次のような仮説 $\mathscr{H}(\omega\,;\Omega)$ の検定に尤度比を用いることは，$\mathscr{H}(\omega\,;\Omega)$ が真のとき，カイ 2 乗分布 $C(n-1)$ を持つ比

$$\frac{(n-1)s^2}{\sigma_0^2}$$

を用いることと同値であることを示せ.　ただし Ω は点 (μ,σ^2) の $\sigma^2\leqslant\sigma_0^2$ なる空間で, ω は $\sigma^2=\sigma_0^2$ なる Ω の部分集合である. よって $\mathscr{H}(\omega\,;\Omega)$ を棄却する大きさ α の棄却域 W_α は

$$\frac{(n-1)s^2}{\sigma_0^2}>\chi_\alpha^2$$

なる s^2 の値の集合である. ここに χ_α^2 はカイ 2 乗分布 $C(n-1)$ の $100\alpha\,\%$ 点である.

13.15　平行回帰直線の仮説検定　$y_{11},\cdots,y_{1n_1}$ は分布 $N(\beta_{10}+\beta_{11}x_{1\xi_1},\sigma^2)$, $\xi_1=1,\cdots,n_1$ を持つ独立な確率変数で,　$y_{21},\cdots,y_{2n_2}$ は分布 $N(\beta_{20}+\beta_{21}p_{2\xi_2},\sigma^2)$, $\xi_2=1,\cdots,n_2$ を持つ独立な確率変数であるとする. $\mathscr{H}(\omega\,;\Omega)$ を次の仮説としよう. Ω は $\sigma^2>0$ なる $(\beta_{10},\beta_{11},\beta_{20},\beta_{21},\sigma^2)$ のユークリッド半空間 R_5 で, ω は $\beta_{11}=\beta_{21}$ なる Ω の部分集合で

ある．$\mathscr{H}$ に対する尤度比は

$$\mathscr{F} = \frac{(S_\omega - S_\Omega)}{S_\omega/(n_1 + n_2 - 4)}$$

に同値であることを示せ．ただし $\mathscr{F}$ は $\mathscr{H}$ が真ならば，スネディッカー分布 $S(1, n_1 + n_2 - 4)$ を持つ．ただし

$$S_\omega = \sum_{p=1}^{2}\sum_{\xi_p=1}^{n_p}[(y_{p\xi_p} - \bar{y}_p) - \hat{\beta}_p(x_{p\xi_p} - \bar{x}_p)]^2$$

$$\bar{y}_p = \frac{1}{n_p}\sum_{\xi_p=1}^{n_p} y_{p\xi_p}, \qquad \bar{x}_p = \frac{1}{n_p}\sum_{\xi_p=1}^{n_p} x_{p\xi_p}, \qquad p = 1, 2,$$

$$\hat{\beta}_p = \frac{a_p}{b_p}, \qquad a_p = \sum_{\xi_p=1}^{n_p}(y_{p\xi_p} - \bar{y}_p)(x_{p\xi_p} - \bar{x}_p), \qquad b_p = \sum_{\xi_p=1}^{n_p}(x_{p\xi_p} - \bar{x}_p)^2$$

で

$$S_\Omega = \sum_{p=1}^{2}\sum_{\xi_p=1}^{n_p}[(y_{p\xi_p} - \bar{y}_p) - \hat{\beta}(x_{p\xi_p} - \bar{x}_p)]^2, \qquad \hat{\beta} = \frac{a_1 + a_2}{b_1 + b_2}$$

である．

13.16　$x_{\xi\eta}$；$\xi = 1, \cdots, r$，$\eta = 1, \cdots, s$ を分布 $N(\mu + \mu_{\xi\cdot} + \mu_{\cdot\eta}, \sigma^2)$ を持つ独立な確率変数とする．ただし $\sum_{\xi=1}^{r}\mu_{\cdot\xi} = 0$，$\sum_{\eta=1}^{s}\mu_{\cdot\eta} = 0$ である．$\mathscr{H}(\omega\,;\Omega)$ を次で定義される仮説とする．すなわち，Ω は $\sum_{\xi}\mu_{\xi\cdot} = \sum_{\eta}\mu_{\cdot\eta} = 0$ なる全実ベクトル $(\mu, \mu_{1\cdot}, \cdots, \mu_{r\cdot}, \mu_{\cdot 1}, \cdots, \mu_{\cdot s}, \sigma^2)$ の集合で，ω は $\mu_{1\cdot} = \cdots = \mu_{r\cdot} = 0$ なる Ω の部分集合とする．$\mathscr{H}$ を検定する尤度比は

$$\mathscr{F} = \frac{S_{\cdot 0}/(r-1)}{S_{\cdot\cdot}/[(r-1)(s-1)]}$$

に同値であることを示せ．ただし $\mathscr{F}$ は $\mathscr{H}$ が真ならば，スネディッカー分布 $S((r-1), (r-1)(s-1))$ を持つ．ただし

$$S_{\cdot 0} = s\sum_{\xi}(\bar{x}_{\xi\cdot} - \bar{x})^2, \qquad S_{\cdot\cdot} = \sum_{\eta}\sum_{\xi}(x_{\xi\eta} - \bar{x}_{\xi\cdot} - \bar{x}_{\cdot\eta} + \bar{x})^2$$

$$\bar{x} = \frac{1}{rs}\sum_{\eta}\sum_{\xi} x_{\xi\eta}, \qquad \bar{x}_{\xi\cdot} = \frac{1}{s}\sum_{\eta} x_{\xi\eta}, \qquad \bar{x}_{\cdot\eta} = \frac{1}{r}\sum_{\xi} x_{\xi\eta}$$

である．

13.17　Luce（1959）の行動選択モデルにおける等確率比検定　この行動モデルでは，2 つの戦略 A_1，A_2 間の選択を行なう場合，A_1 を選択する確率に対する A_2 のその比は第 3 の戦略 A_3 を選択するまで一定であると仮定している．ここでは，実験的な結果に基づいてこの仮説に対する検定を構成する問題を考える．

詳しく述べれば次のようになる．(n_{1i}, n_{2i}, n_{3i})（ただし，$n_{1i} + n_{2i} + n_{3i} = n_i$），$i = 1, \cdots, k$ を 3 項分布

$$\frac{n_i!}{n_{1i}!\,n_{2i}!\,n_{3i}!}p_i^{n_{1i}}q_i^{n_{2i}}r_i^{n_{3i}}$$

を持つ，k 個の独立な確率変数の組とする．ただし，$p_i > 0$，$q_i > 0$，$r_i > 0$ で $p_i + q_i + r_i = 1$ $(i = 1, 2, \cdots, k)$ である．$\mathscr{H}(\omega, \Omega)$ は，Ω が可能なすべての p_i，q_i，r_i，すなわち $p_i + q_i + r_i = 1$，$i = 1, \cdots, k$ を満たす正の数からなる集合で，ω が $q_i/p_i = t$ かつ $p_i(1 + t) + r_i = 1$ なる Ω の部分集合であるような仮説としよう．$\mathscr{H}$ に対する尤度比 λ は

$$\lambda = \frac{\prod_{i=1}^{k}\{(\hat{p}_i)^{n_{1i}+n_{2i}}(\hat{t})^{n_{2i}}[1 - \hat{p}_i(1 + \hat{t})]^{n_{3i}}\}}{\prod_{i=1}^{k}[n_{1i}^{n_{1i}}n_{2i}^{n_{2i}}n_{3i}^{n_{3i}}n^{-n}]}$$

で与えられることを示せ．ここに

$$\hat{t} = \sum_i n_{2i}\Big/\sum_i n_{1i}, \qquad \hat{p}_i = \frac{n_{1i} + n_{2i}}{n_i(1 + \hat{t})}$$

である．さらに，$\mathscr{H}$ が真ならば，$n_1, \cdots, n_k$ が大きいとき，$-2\log\lambda$ は漸近分布として自由度 $k-1$ のカイ 2 乗分布を持つことを示せ．また $\mathscr{H}$ が真のとき，大標本における $\hat{t}$ の分散は

$$t(1 + t)\Big/\Big[\sum_i n_i p_i\Big]$$

で近似されることを示せ．

第14章　ノンパラメトリック仮説検定

第13章では大標本における尤度比検定を中心にパラメトリック統計的仮説検定の理論を述べた．パラメトリック検定の理論においては，許容的な c.d.f. のクラスは $\{F(x,\theta)\colon \theta\in\Omega\}$ なる型，すなわち特定の関数形をした c.d.f. のクラスである．このクラスの要素は，あるパラメータ空間 Ω の実パラメータ θ の値に対応している．もちろん x, θ はどちらも（あるいはどちらか一方が）多次元になってもよい．

一般にノンパラメトリック統計的仮説検定において，許容的な仮説のクラスは，対象となる特定の問題に依存する連続な c.d.f. のクラスであるか，あるいはその部分クラスである．これまでノンパラメトリック検定の一般論は，パラメトリック検定ほど取り扱われてきていなかった．したがって本章では，ノンパラメトリック統計的検定の一般論を打ち立てるというよりも，むしろこの分野における重要な問題を議論していこう．ノンパラメトリック検定の詳細に関する文献に興味を持たれる読者は，Fraser (1957), Kendall (1953) の著作を参照されたい．また Kendall と Sundrum (1953), Moran, Whitfield と Daniels (1950), Scheffé (1943), Wolfowitz (1949), Wilks (1948, 1959 *a*) 等の論文も見られたい．包括的な文献は Savage (1962) により発表されている．

14.1　分位検定

ノンパラメトリック統計的検定の最も簡単な型は，仮説《標本 $(x_1,\cdots,x_n)$ が唯一の p 分位 $\underline{x}_p$ を持つ，すなわち標本が

(14.1.1) $$F(\underline{x}_p)=p$$

なる c.d.f. $F(x)$ の母集団からの標本である》の検定であろう．仮説を明確に述べるた

めに，$\mathscr{C}_{0p}$ を p 分位点として $\underline{x}_{0p}$ を持つ連続な c.d.f. のクラス，$\mathscr{C}_p$ をすべての連続な c.d.f. の集合とする．このとき問題となる仮説を $\mathscr{H}(\mathscr{C}_{0p};\mathscr{C}_p)$ で表わそう．$\mathscr{H}(\mathscr{C}_{0p};\mathscr{C}_p)$ は標本が抽出された母集団の c.d.f. $F(x)$ が $\mathscr{C}_p$ の部分クラス $\mathscr{C}_{0p}$ に属すという仮説である．パラメトリック複合統計的仮説 $\mathscr{H}(\omega;\Omega)$ と区別するために，$\mathscr{H}(\mathscr{C}_{0p};\mathscr{C}_p)$ は基本的にはノンパラメトリック複合統計的仮説ということができる．$\mathscr{H}(\omega;\Omega)$ の場合は，Ω はユークリッド空間の点集合で表わされる元の集合で，ω は Ω の部分集合である．ところが $\mathscr{H}(\mathscr{C}_{0p};\mathscr{C}_p)$ では，許容的な c.d.f. のクラス $\mathscr{C}_p$ はユークリッド空間内の集合への1対1，両連続な対応によって（この集合に）置き換えることができない．$\mathscr{C}_{0p}$ についても同様である．

さて $(x_1,\cdots,x_n)$ を $\mathscr{C}_{0p}$ の c.d.f. からの標本としよう．r を区間 $(-\infty,x_{0p}]$ に落ちる $(x_1,\cdots,x_n)$ の成分の個数とすれば，r は2項分布 $Bi(n,p)$ に従う．$\mathscr{H}(\mathscr{C}_{0p};\mathscr{C}_p)$ に対する直感的に妥当な検定 W_α としては，r を整数からなる2つの集合

(14.1.2) $$\{0,1,\cdots,r_{\frac{1}{2}\alpha}\},\qquad \{r'_{\frac{1}{2}\alpha},\cdots,n\}$$

のどちらか一方に属すという検定が考えられる．ここに $r_{\frac{1}{2}\alpha}$ は

(14.1.3 *a*) $$P(r\leqslant r_{\frac{1}{2}\alpha}\,|\,F\in\mathscr{C}_{0p})\leqslant\frac{1}{2}\alpha$$

なる最大整数，$r'_{\frac{1}{2}\alpha}$ は

(14.1.3 *b*) $$P(r\geqslant r'_{\frac{1}{2}\alpha}\,|\,F\in\mathscr{C}_{0p})\leqslant\frac{1}{2}\alpha$$

なる最小整数である．**9.2.1 a** より，n が大きいとき，r は漸近的に $N(np,\,np(1-p))$ に従って分布する．よって

(14.1.4) $$\lim_{n\to\infty}P(W_\alpha\,|\,F\in\mathscr{C}_{0p})=\alpha.$$

大きな n に対する $r_{\frac{1}{2}\alpha}$, $r'_{\frac{1}{2}\alpha}$ の近似は

(14.1.5) $$\begin{aligned} r_{\frac{1}{2}\alpha}&=np-y_{\frac{1}{2}\alpha}\sqrt{npq}+0(1)\\ r'_{\frac{1}{2}\alpha}&=np+y_{\frac{1}{2}\alpha}\sqrt{npq}+0(1)\end{aligned}$$

となる．ただし $q=1-p$, $y_{\frac{1}{2}\alpha}>0$ で $\Phi(-y_{\frac{1}{2}\alpha})=\frac{1}{2}\alpha$, $\Phi(x)$ は $N(0,1)$ の c.d.f. である．

さて W_α の一致性を調べよう．$F(x)$ が $\mathscr{C}_p-\mathscr{C}_{0p}$ の任意の要素ならば，r は2項分布 $Bi(n,p')$, $(p'\neq p)$ を持つ．$F\in\mathscr{C}_{0p}$ ならば，$n\longrightarrow\infty$ のとき r/n は定数 p に確率収束（区間 $(r_{\frac{1}{2}\alpha}/n,r'_{\frac{1}{2}\alpha}/n)$ は点 p に収束）し，$F\in\mathscr{C}_p-\mathscr{C}_{0p}$ ならば，同じく p'

$(\neq p)$ に確率収束する．よって

$$\lim_{n\to\infty} P(W_\alpha | F \in \mathscr{C}_p - \mathscr{C}_{0p}) = 1 \tag{14.1.6}$$

は明らかである．ゆえに

14.1.1 r が (14.1.2) での整数からなる2つの集合のどちらか一方に属すという検定 W_α は仮説 $\mathscr{H}(\mathscr{C}_{0p}; \mathscr{C}_p)$ の一致検定である．

$r \in \{0, 1, \cdots, r_\alpha\}$ に対する検定 W_α^L は $\underline{x}_{0p}$ より大きい p 分位を持つ $\mathscr{C}_p$ の c.d.f. に対する $\mathscr{C}_{0p}$ の任意の c.d.f. の一致検定となる**左片側検定**である．同様に，$r \in \{r'_\alpha, \cdots, n\}$ に対する検定 W_α^U は $\underline{x}_{0p}$ より小さい p 分位を持つ $\mathscr{C}_p$ の c.d.f. に対する $\mathscr{C}_{0p}$ の任意の c.d.f. の一致検定となる**右片側検定**である．

注意　$p = 0.5$ のとき，検定 W_α は**符号検定**といわれる．このとき，仮説 $\mathscr{H}(\mathscr{C}_{0p}; \mathscr{C}_p)$ は《$(x_1, \cdots, x_n)$ がメディアンとして $\underline{x}_{0.5}$ を持つ c.d.f. $F(x)$ からの標本である》という仮説に帰着する．実際問題で重要なのは，標本成分 $(x_1, \cdots, x_n)$ 自身が確率変数の独立な対 $(u_1, v_1; u_2, v_2; \cdots; u_n, v_n)$ の差，すなわち $x_\xi = u_\xi - v_\xi$, $\xi = 1, \cdots, n$ になるときである．このような例題においては，通常メディアン $\underline{x}_{0.5}$ は0である．符号検定に関する全般的な議論については Dixon と Mood (1946) を参照されたい．メディアンに関する詳細な検定は Walsh (1949) に展開されている．

14.2 ノンパラメトリック単純統計的仮説

(a) 基本的な注意

基本的なノンパラメトリック統計的仮説の1つに次が考えられる．標本 $(x_1, \cdots, x_n)$ はある連続な c.d.f. $F(x)$ からとする．この標本がはたして特定の連続な c.d.f. $F_0(x)$ からの標本であると“理論的”に判定することができるだろうか？ $\mathscr{C}$ を連続な c.d.f. のクラスとして便宜上この仮説を $\mathscr{H}(F_0; \mathscr{C})$ で表わし，**ノンパラメトリック単純統計的仮説**という．

理想的には $\mathscr{H}(F_0; \mathscr{C})$ に対する一致検定，すなわち無限に大きな標本の場合において F_0 と $\mathscr{C}$ の F_0 と異なる任意の c.d.f. とを区別できるような検定方式を開発したい．しかしこれは大変難かしい．多くの場合，$\mathscr{C}$ の種々の部分クラスに含まれる c.d.f. に対して，F_0 の一致検定を構成することができる．

$\mathscr{H}(F_0; \mathscr{C})$ の検定方式に関する問題について3つのアプローチが考えられる．最初は，ピヤソンのカイ2乗検定に基づくノンパラメトリック複合仮説による接近法，次は**空細胞検定**といわれるもので，絶対連続な c.d.f. のクラス $\mathscr{A}$ の c.d.f. に対する単純ノンパラメトリック接近法，最後は 11.6, 11.7 節で議論された信頼領域に基づく，連続な c.d.f. のクラスの c.d.f. に対するノンパラメトリックな処理方法である．

(b) ピヤソンのカイ2乗検定に基づくノンパラメトリック複合仮説

$\underline{x}_{i/m}$ を c.d.f. $F_0(x)$ の i/m 分位とする．すなわち $F_0(\underline{x}_{i/m}) = i/m$, $i = 1, \cdots, m$. $\mathscr{C}_0$ をこれらと同じ分位を持つ $\mathscr{C}$ の c.d.f. の部分クラスとする．I_i $(i = 1, \cdots, m)$ は区間 $(\underline{x}_{(i-1)/m}, \underline{x}_{i/m}]$ である．ただし $\underline{x}_0 = -\infty$, $\underline{x}_1 = +\infty$ とする．r_i は I_i に落ちる標本成分の個数である．標本が $\mathscr{C}_0$ の任意の c.d.f. からであれば，細胞頻度 $(r_1, \cdots, r_m)$ は $r_1 + \cdots + r_m = n$ を満たす m 次元確率変数で

$$p(r_1, \cdots, r_m) = \frac{n!}{r_1! \cdots r_m!} \left(\frac{1}{m}\right)^n \tag{14.2.1}$$

なる p.f. を持つ $(m-1)$ 次元多項分布 $M(n; 1/m, \cdots, 1/m)$ に従う．

すべての n に対して m を固定して考えれば，後で表現されるある種の犠牲を払えば，$\mathscr{H}(F_0; \mathscr{C})$ の検定問題は次の比較的簡単な問題に置き換えられる．すなわちピヤソンのカイ2乗検定基準を用いて，ノンパラメトリック複合仮説を検定する問題である．この検定は後で見るように，ノンパラメトリック複合仮説 $\mathscr{H}(\mathscr{C}_0; \mathscr{C})$，すなわち，対立仮説 $F(x) \in \mathscr{C} - \mathscr{C}_0$ に対する任意の仮説 $F(x) \in \mathscr{C}_0$ の一致検定である．そしてこれは多くの実際的な統計の目的に対しては $\mathscr{H}(F_0; \mathscr{C})$ のより細分化されたノンパラメトリック検定のかわりに用いられる．

m を固定して

$$Q_n = \frac{m}{n} \sum_{i=1}^{m} \left(r_i - \frac{n}{m}\right)^2 \tag{14.2.2}$$

とし，$W_{1\alpha}$ を $(r_1, \cdots, r_m)$ が

$$Q_n > \chi^2_{\alpha, n} \tag{14.2.3}$$

を満たす検定とする．ただし $\chi^2_{\alpha,n}$ は，標本が $\mathscr{C}_0$ の c.d.f. からのとき，各 n について $W_{1\alpha}$ の大きさが α に限りなく近づくように選ばれるとする．

9.3.2a より

(14.2.4)
$$\lim_{n\to\infty} P(W_{1\alpha}|F \in \mathscr{C}_0) = \int_{\chi^2_\alpha}^{\infty} dF_{m-1}(\chi^2) = \alpha$$

が成り立つ．ただし $dF_{m-1}(\chi^2)$ は (7.8.1) で与えられる自由度 $m-1$ のカイ2乗分布の p.e. である．

さて $\mathscr{H}(\mathscr{C}_0\,;\mathscr{C})$ の検定に対する $W_{1\alpha}$ の検定力を考察しよう．つまり $P(W_{1\alpha}|F \in \mathscr{C} - \mathscr{C}_0)$ の値を調べる．p_i を $\mathscr{C} - \mathscr{C}_0$ のある c.d.f. $F_1(x)$ から計算された I_i の確率とすれば

(14.2.5)
$$p_i = \int_{I_i} dF_1(x), \qquad i = 1, \cdots, m.$$

もちろん点 $(p_1, \cdots, p_m)$ は $(1/m, \cdots, 1/m)$ と異なる．

9.1.1 の多次元への拡張より $(x_1, \cdots, x_n)$ が c.d.f. $F_1(x)$ からの標本ならば，確率変数 $(r_1/n, \cdots, r_m/n)$ は $n \longrightarrow \infty$ のとき点 $(p_1, \cdots, p_m)$ に確率収束する．さて与えられた n に対して

(14.2.6)
$$\sum_{i=1}^{m}\left(\frac{r_i}{n} - p_i\right)^2 < D^2$$

なる $(x_1, \cdots, x_n)$ の標本空間内の集合 E_n を考えよう．ただし D^2 は2点 $(p_1, \cdots, p_m)$, $(1/m, \cdots, 1/m)$ 間の（ユークリッド）距離の平方より小さい正の数である．このとき $W_{1\alpha}$ が

(14.2.7)
$$\sum_{i=1}^{m}\left(\frac{r_i}{n} - \frac{1}{m}\right)^2 > \frac{\chi^2_{\alpha,n}}{mn}$$

を満たす標本点からなる（ただし $\lim_{n\to\infty} \chi^2_{\alpha,n} = \chi^2_\alpha$）ことより，ある n_1 より大きい n に対して

$$E_n \subset W_{1\alpha}$$

になる．したがって，$n > n_1$ ならば

(14.2.8)
$$P(E_n|F_1) \leqslant P(W_{1\alpha}|F_1).$$

しかし $n \longrightarrow \infty$ のとき $(r_1/n, \cdots, r_m/n)$ は $(p_1, \cdots, p_m)$ に確率収束するから，任意の $\delta > 0$ に対して n_2 が存在し，$n > n_2$ ならば

$$P(E_n|F_1) > 1 - \delta.$$

よって $n > \max(n_1, n_2)$ のとき

$$P(W_{1\alpha}|F_1) \geqslant 1 - \delta$$

すなわち

(14.2.9)
$$\lim_{n\to\infty} P(W_{1\alpha}|F_1) = 1.$$

このように $W_{1\alpha}$ は $\mathscr{C} - \mathscr{C}_0$ に属す任意の c.d.f. に対して $F(x)$ が $\mathscr{C}_0$ に属す仮説 $\mathscr{H}(\mathscr{C}_0\,;\mathscr{C})$ の一致検定である．もちろん $\mathscr{C}_0$ の他の c.d.f. に対する $F_0(x)$ の検定については一致性はない．

要約すると次のようになる．

14.2.1　$\mathscr{C}_0$ を分位 $\underline{x}_{i/m}$, $i = 1, \cdots, m-1$ を持つ（$F_0(x)$ を含む）連続な c.d.f. のクラス，$I_i = (\underline{x}_{(i-1)/m}, \underline{x}_{i/m}]$, $(\underline{x}_0 = -\infty, \underline{x}_1 = +\infty)$ をこれらの分位より定まる区間とする．$r_1, \cdots, r_m$ をそれぞれ区間 $I_1, \cdots, I_m$ に落ちる標本 $(x_1, \cdots, x_n)$ の成分の個数とする．$W_{1\alpha}$ は

$$\frac{m}{n}\sum_{i=1}^{m}\left(r_i - \frac{n}{m}\right)^2 > \chi^2_{\alpha,n}$$

なる $(x_1, \cdots, x_n)$ の標本空間の点からなる検定である．ただし $\chi^2_{\alpha,n}$ は任意の $F \in \mathscr{C}_0$ に対して，検定 $W_{1\alpha}$ の大きさが限りなく α に近づくように選ぶ．このとき任意の $F \in \mathscr{C}_0$ に対して (14.2.4) が成り立つ．検定 $W_{1\alpha}$ は任意の $F \in \mathscr{C} - \mathscr{C}_0$ に対する任意の $F \in \mathscr{C}_0$ の一致検定である．しかし任意の $F \in \mathscr{C}_0$ に対して (14.2.4) が成り立つ．よって $W_{1\alpha}$ は $\mathscr{C}_0$ の任意の c.d.f. に対する F_0 の一致検定でない．

(c)　空 細 胞 検 定

上で考察したように，ノンパラメトリック単純仮説 $\mathscr{H}(F_0\,;\mathscr{C})$ の検定問題に対する接近方法は $W_{1\alpha}$ が $\mathscr{C}_0$ の対立する c.d.f. に対する F_0 の一致検定でないという大きな欠点をもっていた．もちろんこのクラスは $n \longrightarrow \infty$ のとき m を大きい値に固定すれば弱められる．しかし $n \longrightarrow \infty$ としても m を固定する限り，この検定は本質的には F_0 と $\mathscr{C}_0$ の他のすべての要素とを区別できないノンパラメトリック複合統計的検定である．したがって次のような問題が考えられる．m, n をともに限りなく大きくすることによって，F_0 と $\mathscr{C}_0$ の"ほとんどすべて"の c.d.f. とを区別するような検定方式を開発できないだろうか？ F_0 と $\mathscr{C}$ の部分クラス $\mathscr{A}$ の c.d.f. とを区別する簡単な検定を考えよう．ここに $\mathscr{A}$ は絶対連続な c.d.f. のクラス，すなわち導関数 (p.d.f.) $\{f(x)\}$ を持つ c.d.f.

のクラス $\{F(x)\}$ である．もちろん $F_0(x)$ も $\mathscr{A}$ に属すとする．

$s_0, s_1, \cdots, s_n$ を標本 $(x_1, \cdots, x_n)$ より定まる**細胞頻度数**，すなわち標本成分をそれぞれ $0, 1, \cdots, n$ 個含む区間 I_i，$i=1, \cdots, m$ の個数とする．このとき標本が $\mathscr{C}_0$ の任意の c.d.f. からであれば，（退化）確率変数 $(s_0, s_1, \cdots, s_n)$ の p.f. は (14.2.1) より容易に得られ，次式で表わされる．

(14.2.10)
$$p(s_0, s_1, \cdots, s_n) = \frac{m!\,n!}{m^n(0!)^{s_0}(1!)^{s_1}\cdots(n!)^{s_n}s_0!\,s!\cdots s_n!}.$$

もちろん $s_0, s_1, \cdots, s_n$ は2つの条件式

(14.2.11)
$$s_0+s_1+\cdots+s_n=m, \qquad s_1+2s_2+\cdots+ns_n=n$$

を満たす非負整数である．

(6.1.4) および (6.1.7) と同様にすると，確率変数 $(s_0, s_1, \cdots, s_n)$ の成分の一般階乗モーメントが次のように求められる．

(14.2.12)
$$\mathscr{E}(s_0^{[g_0]}s_1^{[g_1]}\cdots s_n^{[g_n]})$$
$$=\frac{m!\,n!(m-g_0-g_1-\cdots-g_n)^{n-g_1-2g_2-\cdots-ng_n}}{m^n(2!)^{g_2}\cdots(n!)^{g_n}(m-g_0-g_1-\cdots-g_n)!\,(n-g_1-2g_2-\cdots-ng_n)!}.$$

ただし非負整数 $g_0, \cdots, g_n$ は条件 $m-g_0-g_1-\cdots-g_n \geqslant 0$ および $n-g_1-2g_2-\cdots-ng_n \geqslant 0$ を満たさなければならない．上式より $s_0, s_1, \cdots, s_n$ の平均，分散，共分散および他のモーメントが求められる．

David (1950) が提唱した**空細胞検定**は，標本成分を1つも含まない区間 $I_1, \cdots, I_m$ の個数 s_0 に基づいている．(14.2.12) から s_0 の平均と分散は次のようになる．

(14.2.13)
$$\mu(s_0)=m\left(1-\frac{1}{m}\right)^n$$
$$\sigma^2(s_0)=m(m-1)\left(1-\frac{2}{m}\right)^n+m\left(1-\frac{1}{m}\right)^n-m^2\left(1-\frac{1}{m}\right)^{2n}.$$

実際に (14.2.10) の構造を調べると，$p(s_0, s_1, \cdots, s_n)$ は

(14.2.14)
$$\frac{n!}{m^n}\left(u_0+\frac{u_1v}{1!}+\frac{u_2v^2}{2!}+\cdots+\frac{u_nv^n}{n!}\right)^m$$

の形式的展開における $u_0^{s_0}u_1^{s_1}\cdots u_n^{s_n}v^n$ の係数であることがわかる．よって s_0 の p.f. $p(s_0)$ は (14.2.14) で $u_1=\cdots=u_n=1$ と置いたときの式

(14.2.15)
$$\frac{n!}{m^n}\left(u_0+\frac{v}{1!}+\frac{v^2}{2!}+\cdots+\frac{v^n}{n!}\right)^m$$

の $u_0^{s_0}v^n$ の係数をとることによって得られる．

しかし (14.2.15) における $u_0^{s_0}v^n$ の係数は

(14.2.16)
$$\frac{n!}{m^n}(u_0+e^v-1)^m$$

の同じ係数に等しい．また

(14.2.17)
$$\varphi(v)=\frac{n!}{m^n}\binom{m}{s_0}(e^v-1)^{m-s_0}$$

における v^n の係数にも等しい．ところが $\varphi(v)$ における v^n の係数は

(14.2.18)
$$\frac{1}{n!}\left[\frac{d^n\varphi(v)}{dv^n}\right]_{v=0}$$

で与えられる．この微分演算を行ない，$v=0$ とすれば

(14.2.19)
$$p(s_0)=\frac{m!}{m^ns_0!}\sum_{i=0}^{m-s_0}\frac{(-1)^i(m-s_0-i)^n}{i!\,(m-s_0-i)!}$$

を得る．ただし $s_0=k, k+1, \cdots, m-1$；$k=\max(0, m-n)$.

m と n の値が大きい場合は David (1950) による次の結果がただちに成立する．

14.2.2　$(x_1, \cdots, x_n)$ が連続な c.d.f. $F_0(x)$ からの標本ならば，$n/m \longrightarrow \rho>0$ となるように $m, n \longrightarrow \infty$ とするとき，s_0 は漸近的に $N(me^{-\rho}, m[e^{-\rho}-e^{-2\rho}(1+\rho)])$ に従って分布する．

$W_{2\alpha}$ を $(x_1, \cdots, x_n)$ の標本空間において $s_0>s_{0\alpha}$ となる検定とする．ただし $s_{0\alpha}$ は $\sum_{s_0=s_{0\alpha}}^{m-1}p(s_0) \leqslant \alpha$ を満たす最小整数である．**14.2.2** より，$n=\rho m+O(1)$ なる大きな m, n に対して

(14.2.20)
$$s_{0\alpha}=m\left[e^{-\rho}+\frac{y_\alpha}{\sqrt{m}}(e^{-\rho}-e^{-2\rho}(1+\rho))^{1/2}+O\left(\frac{1}{m}\right)\right].$$

そして $n/m \longrightarrow \rho>0$ を満たしながら $m, n \longrightarrow \infty$ とすれば

(14.2.21)
$$\lim P(W_{2\alpha}|F_0)=\frac{1}{\sqrt{2\pi}}\int_{y_\alpha}^{\infty}e^{-\frac{1}{2}t^2}dt=\alpha$$

となる．

標本が c.d.f. $F_0(x)$ を持つ分布と異なる $\mathscr{C}$ の任意の分布からならば，s_0 の分布が s_0 軸の右側に偏在しがちであることは直感的に明らかである．よってこの $W_{2\alpha}$ の選択は妥当である．

$W_{2\alpha}$ の妥当性を形式的に調べるために絶対連続な c.d.f. のクラス $\mathscr{A}$ の対立する要素（F_0 と異なる）に対する絶対連続な F_0 の検定の一致性を考えよう．$f_0(x)$ が $F_0(x)$

の導関数，$f_1(x)$ が $F_0(x)$ と異なる $\mathscr{A}$ の任意の c. d. f. $F_1(x)$ の導関数であれば，$f_0(x)$ と $f_1(x)$ はこのどちらかで計算された確率正の集合上では異なる．$\mathscr{A}^*$ を比 $f_1(x)/f_0(x)$ の $F_0(x)$, $F_1(x)$ のどちらによるモーメントもすべて存在するような $\mathscr{A}$ の部分クラスとする．$\mathscr{A}^*$ は $F_0(x)$ を含む．ここで次の結果を得る．

14. 2. 3　　検定 $W_{2\alpha}$ は $\mathscr{A}^*$ の任意の $F_1(x)$ $(\neq F_0(x))$ に対する $F_0(x)$ の検定に対して一致性を持つ．すなわち $n/m \longrightarrow \rho > 0$ として，$m, n \longrightarrow \infty$ ならば

$$\lim P(W_{2\alpha} \mid F_1 \in \mathscr{A}^* - F_0) = 1.$$

14. 2. 3 の証明には，標本 $(x_1, \cdots, x_n)$ が $\mathscr{A}^* - F_0$ の任意の $F_1(x)$ からのとき，s_0/m は $n/m \longrightarrow \rho > 0$, $m, n \longrightarrow \infty$ のとき $e^{-\rho}$ を超える定数に確率収束することを示せば十分である．

z_i $(i = 1, \cdots, m)$ を I_i が標本成分を1つも含まないとき1，その他のとき0を値に持つ確率変数とする．このとき

$$s_0 = z_1 + \cdots + z_m.$$

$\mathscr{E}(s_0/m \mid F_1 \in \mathscr{A}^* - F_0)$ を $\mathscr{E}(s_0/m \mid F_1)$ で表わせば

$$\mathscr{E}\left(\frac{s_0}{m}\middle| F_1\right) = \frac{1}{m}[\mathscr{E}(z_1) + \cdots + \mathscr{E}(z_m)].$$

しかし

$$\mathscr{E}\left(\frac{z_i}{m}\right) = \frac{1}{m}(1 - p_i)^n, \qquad i = 1, \cdots, m.$$

ただし p_i は $f_1(x)$ で計算した I_i の確率である．したがって

(14. 2. 22)
$$\mathscr{E}\left(\frac{s_0}{m}\middle| F_1\right) = \frac{1}{m}\sum_{i=1}^{m}(1 - p_i)^n$$
$$= 1 - \frac{n}{m} + \frac{n(n-1)}{2!\,m}\sum_i p_i^2 - \frac{n(n-1)(n-2)}{3!\,m}\sum_i p_i^3 + \cdots$$
$$+ (-1)^k\frac{n(n-1)\cdots(n-k+1)}{k!\,m}\sum_i p_i^k + \cdots + \frac{(-1)^n}{m}\sum_i p_i^n.$$

上式の一般項は次のようになる．

$$(-1)^k\frac{1}{k!}\left(\frac{n}{m}\right)\left(\frac{n-1}{m}\right)\cdots\left(\frac{n-k+1}{m}\right)\left\{\sum_{i=2}^{m-1}\left[\left(\frac{p_i}{\Delta_i}\right)\Big/\left(\frac{\left(\frac{1}{m}\right)}{\Delta_i}\right)\right]^k\frac{1}{m}\right\} + \delta_{m,n}$$

ただし Δ_i は I_i の長さで

$$\delta_{m,n} = (-1)^k\frac{1}{k!}\left(\frac{n}{m}\right)\left(\frac{n-1}{m}\right)\cdots\left(\frac{n-k+1}{m}\right)\left[\left(\frac{p_1}{1/m}\right)^k + \left(\frac{p_m}{1/m}\right)^k\right]\frac{1}{m}.$$

$n/m \longrightarrow \rho > 0$, $m, n \longrightarrow \infty$ として極限をとれば，$f_1(x)/f_0(x)$ の $F_0(x)$, $F_1(x)$ のどちらによるモーメントもすべて有限であることを仮定しているので $\delta_{m,n} \longrightarrow 0$ となり，(14. 2. 22) の極限は

$$\sum_{k=0}^{\infty}(-1)^k\frac{\rho^k}{k!}\int_{-\infty}^{\infty}[h(x)]^k\,dF_0(x)$$

になる．ただし $h(x) = f_1(x)/f_0(x)$．よって $n/m \longrightarrow \rho > 0$ として，$m, n \longrightarrow \infty$ ならば

(14. 2. 23)
$$\lim \mathscr{E}\left(\frac{s_0}{m}\middle| F_1\right) = \int_{-\infty}^{\infty} e^{-\rho h(x)}\,dF_0(x)$$

が得られる．

一般に $g(x)$ が非負確率変数で，その平均が有限のとき

(14. 2. 24)
$$\mathscr{E}(g(x)) \geqslant e^{\mathscr{E}(\log g(x))}$$

で，$g(x)$ が確率1で定数のときのみ等号が成り立つことを利用すれば

(14. 2. 25)
$$\int_{-\infty}^{\infty} e^{-\rho h(x)}\,dF_0(x) \geqslant \exp\left[-\int_{-\infty}^{\infty}\rho h(x)\,dF_0(x)\right] = e^{-\rho}.$$

等号は確率1で $h(x) \equiv C$ のとき，かつこのときに限り成立する．しかしこれは確率1で $f_0(x) \equiv f_1(x)$ を意味する．よって $F_0(x) \equiv F_1(x)$ である．このように $f_0(x)$ と $f_1(x)$ が確率正の集合上で異なれば，すなわち $F_1 \in \mathscr{A}^* - F_0$ ならば

(14. 2. 26)
$$\lim \mathscr{E}\left(\frac{s_0}{m}\middle| F_1\right) > e^{-\rho}$$

が成立する．

14. 2. 3 の議論を終えるには，任意の n に対して，$m \longrightarrow \infty$ のとき $\sigma^2(s_0/m \mid F_1) \longrightarrow 0$ を示せばよい．$(z_1, \cdots, z_m)$ の共分散行列を $\|\sigma_{ij}\|$ で表わせば

(14. 2. 27)
$$\sigma^2\left(\frac{s_0}{m}\middle| F_1\right) = \frac{1}{m^2}\sum_{i,j=1}^{m}\sigma_{ij} = \frac{1}{m^2}\sum_{i=1}^{m}(1 - p_i)^n[1 - (1 - p_i)^n]$$
$$+ \frac{1}{m^2}\sum_{i \neq j=1}^{m}[(1 - p_i - p_j)^n - (1 - p_i)^n(1 - p_j)^n].$$

$(1 - p_i - p_j)^n \leqslant (1 - p_i)^n(1 - p_j)^n$ だから

$$\sigma^2\left(\frac{s_0}{m}\middle| F_1\right) \leqslant \frac{1}{m^2}\sum_{i=1}^{m}(1 - p_i)^n[1 - (1 - p_i)^n].$$

ところが任意の n について $(1-p_i)^n[1-(1-p_i)^n] \leqslant \frac{1}{4}$, $i=1,\cdots,m$. よって

$$\sigma^2\left(\frac{s_0}{m}\middle|F_1\right) \leqslant \frac{1}{4m}.$$

このことより, $m \longrightarrow \infty$ ならば $\sigma^2(s_0/m|F_1) \longrightarrow 0$ が成り立つ. $\mathscr{E}(s_0/m|F_1) > \mathscr{E}(s_0/m|F_0) = e^{-\rho}$ および $m \longrightarrow \infty$ のとき $\sigma^2(s_0/m|F_1) \longrightarrow 0$ から, $n/m \longrightarrow \rho > 0$, $m, n \longrightarrow \infty$ ならば

$$\lim P(W_{2\alpha}|F_1 \in \mathscr{A}^* - F_0) = 1$$

を得る. これで **14.2.3** の議論を終わる.

事実, Okamoto (1952) および Kitabatake (1958) により, 標本 $(x_1, \cdots, x_n)$ が $\mathscr{A}^*$ の c.d.f. $F_1(x)$ を持つ任意の分布からのとき, $n = \rho m + O(1)$, $\rho > 0$ なる大きな m, n に対して s_0 の漸近分布は

$$N\left(m\int_0^1 e^{-\rho h(x)}\,dF_0(x), mK^2\right)$$

であることが示されている. ただし

(14.2.28)
$$K^2 = \int_{-\infty}^{\infty}[e^{-\rho h(x)} - e^{-2\rho h(x)}]\,dF_0(x) - \rho\left[\int_{-\infty}^{\infty} e^{-\rho h(x)}h(x)\,dF_0(x)\right]^2.$$

(d)　ノンパラメトリック検定としての信頼水準

11.6, 11.7 節において, 分布からの標本 $(x_1, \cdots, x_n)$ による分布の c.d.f. $F(x)$ の推定問題を議論した. これらの推定問題では, (11.6.2) で定義された経験 c.d.f. $F_n(x)$ が基本的な統計関数である. $\mathscr{C}_1$ を連続な c.d.f. クラス $\mathscr{C}$ のある部分クラスとしたとき, $\mathscr{H}(F_0;\mathscr{C}_1)$ の検定方式に関する問題を扱う場合にも, $F_n(x)$ 自身が大切な役割を果たすであろう.

Anderson と Darling (1952), Cramér (1928), Kimball (1947), Kolmogorov (1933 *b*), Malmquist (1954), Sherman (1950), Smirnov (1939 *b*), von Mises (1931) 等は $F_n(x)$ に基づく検定を考察している. ここでは Birnbaum (1953 *b*), Darling (1957) による検定を比較検討しよう. これらの種々の検定のうち, 11.6, 11.7 節では信頼水準だけを論じてきた.

11.6 節に戻って, $W_{3\alpha}$ をすべての x については

(14.2.29)
$$F_0(x) \leqslant F_n(x) + d$$

が成立しないような $(x_1, \cdots, x_n)$ の標本空間における事象とする. d を $P(W_{3\alpha}|F_0) = \alpha$ を満たすように選べば, $W_{3\alpha}$ は大きさ α の検定である. ここでは検出力を調べよう. もっと厳密にいえば, $F_1(x)$ が $F_0(x)$ と異なる $\mathscr{C}$ の要素で次に述べるような比較的ゆるい条件を満たすとき, Birnbaum (1953 *a*) による結果を用いて, $P(W_{3\alpha}|F_1)$ の下限を求めることである. さらにこの条件の下では, $W_{3\alpha}$ は $\mathscr{C}$ の広い部分クラスの c.d.f. に対する F_0 の一致検定であることも示される.

11.6.1 より, $(x_1, \cdots, x_n)$ が $F_0(x)$ からの標本のとき

(14.2.30)
$$P(F_0(x) \leqslant F_n(x) + d,\ \text{すべての } x \text{ に対して} \mid F_0) = P_n(d)$$

すなわち

(14.2.31)
$$P(W_{3\alpha}|F_0) = 1 - P_n(d)$$

と書ける. ただし $P_n(d)$ は公式 (11.6.5) で与えられている. 与えられた有意水準 α と標本の大きさ n に対して (11.6.5) から d (これを $d'(n,\alpha)$, あるいは簡単に d' で表わす) の値を求めると, $P_n(d') = 1-\alpha$ より, 大きさ α の検定 $W_{3\alpha}$ が構成される. すなわち

(14.2.32)
$$P(W_{3\alpha}|F_0) = \alpha.$$

さて $F_1(x)$ を $F_0(x)$ と異なる $\mathscr{C}$ の c.d.f. とする. 対立仮説 $F_1(x)$ に対する検定 $W_{3\alpha}$ の検出力 $P(W_{3\alpha}|F_1)$ を求めよう. ただし

(14.2.33)
$$1 - P(W_{3\alpha}|F_1) = P(F_0(x) \leqslant F_n(x) + d',\ \text{すべての } x \text{ に対して} \mid F_1).$$

しかし $F_0(x) \leqslant F_n(x) + d'$ がすべての x について成立するための必要十分条件は, この不等式が $F_n(x)$ の跳びで成立することである. すなわち必要十分条件は

(14.2.34)
$$F_0(x_{(\xi)}) \leqslant \frac{\xi-1}{n} + d', \qquad \xi = 1, \cdots, n.$$

ただし $(x_{(1)}, \cdots, x_{(n)})$ は標本順序統計量である. しかし (14.2.34) が成り立つための必要十分条件は

(14.2.35)
$$x_{(\xi)} \leqslant F_0^{-1}\left(\frac{\xi-1}{n} + d'\right), \qquad \xi = 1, \cdots, n.$$

ここに $F_0^{-1}(y)$ は $F_0(x)$ の逆関数である. さらに, (14.2.35) が成立するための必要十分条件は

(14.2.36)
$$F_1(x_{(\xi)}) \leqslant F_1\left[F_0^{-1}\left(\frac{\xi-1}{n} + d'\right)\right], \qquad \xi = 1, \cdots, n$$

が成立することである.

(14.2.37)
$$F_1(F_0^{-1}(z)) = G(z)$$

とし，確率変数の変換 $y_{(\xi)}=F_1(x_{(\xi)})$, $\xi=1,\cdots,n$ を行なう．ところが (8.7.2) より $(y_{(1)},\cdots,y_{(n)})$ の p.e. は

(14.2.38)
$$n!\,dy_{(1)}\cdots dy_{(n)}.$$

したがって (14.2.36) が成り立つ確率，すなわち (14.2.33) の値は

(14.2.39)
$$n!\int_0^{G(d')}\int_{y(1)}^{G(1/n+d')}\cdots\int_{y(n-1)}^{G((n-1)/n+d')}dy_{(n)}\,dy_{(n-1)}\cdots dy_{(1)}$$

で与えられる．

さて

(14.2.40)
$$\underset{-\infty<x<+\infty}{\text{l.u.b.}}\;(F_0(x)-F_1(x))=\delta$$

を仮定し

(14.2.41)
$$F_0(x')-F_1(x')=\delta$$

となる x の値を x' としよう．k が $n(F_0(x')-d')$ 以下の最大整数であれば

(14.2.42)
$$G\left(\frac{\xi-1}{n}+d'\right)\leqslant F_0(x')-\delta,\qquad \xi\leqslant k+1.$$

さらに

(14.2.43)
$$G\left(\frac{\xi-1}{n}+d'\right)\leqslant 1,\qquad \xi=k+2,\cdots,n.$$

(14.2.39) における積分の上限 $G[(\xi-1)/n+d']$, $\xi=1,\cdots,n$ を (14.2.42), (14.2.43) の右辺の値で置き換えても積分の値は減少しない．この新しい上限を持つ積分は帰納的に評価され，値

(14.2.44)
$$1-\sum_{\xi=0}^{k}\binom{n}{\xi}Q^{\xi}(1-Q)^{n-\xi}$$

を持つ．ただし

(14.2.45)
$$Q=F_0(x')-\delta.$$

k は $n(F_0(x')-d')$ 以下，すなわち $n(Q+\delta-d')$ 以下の最大整数だから，次の検出力の下限が得られる．

(14.2.46)
$$P(W_{3\alpha}|F_1)\geqslant\sum_{\xi=0}^{k}\binom{n}{\xi}Q^{\xi}(1-Q)^{n-\xi}.$$

要約すると次のようになる．

14.2.4　標本の大きさ n に対して，$W_{3\alpha}$ を (14.2.29) が成り立たないような標本空間の点集合とする．$F_1(x)$ は (14.2.40) が成立するような $\mathscr{C}$ の要素とする．このとき検定 $W_{3\alpha}$ の検出力 $P(W_{3\alpha}|F_1)$ は (14.2.46) で示される下限を持つ．

大標本の場合は，**9.2.1 a** から (14.2.46) の右辺は近似的に

(14.2.47)
$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}e^{-\frac{1}{2}t^2}\,dt$$

になる．ただし

(14.2.48)
$$y=\frac{\sqrt{n}\,(\delta-d')}{\sqrt{Q(1-Q)}}.$$

さて d' は n と α の関数で $\lim\limits_{n\to\infty}d'=0$ となる．よって $\delta>0$ ならば $n\longrightarrow\infty$ のとき積分 (14.2.47) の極限は 1 になる．ゆえに $\delta>0$ のとき

(14.2.49)
$$\lim_{n\to\infty}P(W_{3\alpha}|F_1)=1$$

である．したがって次の結果が得られる．

14.2.5　$\mathscr{C}_1$ を (14.2.40) の $\delta>0$ となる $\mathscr{C}$ の部分クラスとする．このとき $W_{3\alpha}$ は，ノンパラメトリック単純仮説 $\mathscr{H}(F_0,\mathscr{C}_1\cup F_0)$ に対する一致検定である．

この結果の概要は次の事柄を意味している．$F_0(x)$ のグラフが $F_1(x)$ のグラフより，どこかで上にあるとき $(x_1,\cdots,x_n)$ が $F_1(x)$ からの標本ならば，$n\longrightarrow\infty$ のとき $F_n(x)+d'$ のグラフがどこかで $F_0(x)$ の下に位置する確率は 1 に近づく．

(11.6.1) より，$W'_{3\alpha}$ をすべての x に対しては

(14.2.29 a)
$$F_0(x)\geqslant F_n(x)-d'$$

が成立しないような標本空間の点集合で定義された検定とすれば，$W'_{3\alpha}$ の大きさは α になる．さらに $\mathscr{C}'_1$ が

(14.2.40 a)
$$\underset{-\infty<x<+\infty}{\text{g.l.b.}}\;(F_0(x)-F_1(x))=\delta'$$

で $\delta'<0$ となる $\mathscr{C}$ の部分クラスならば，**14.2.5** と同様に，$W'_{3\alpha}$ は $\mathscr{H}(F_0;\mathscr{C}'_1\cup F_0)$ に対する一致検定である．読者には $W'_{3\alpha}$ に関する **14.2.4**, **14.2.5** に対応する定理の公式化を残しておこう．

さて，**14.2.5** および $W'_{3\alpha}$ は $\mathscr{C}'_1$ の c.d.f. に対する $F_0(x)$ の一致検定であることに注意しよう．このとき，$W''_{3\alpha}$ を棄却域としてすべての x に対しては

(14.2.29 b)
$$|F_n(x)-F_0(x)|\leqslant d''$$

が成立しないような点集合を持つ検定とし，d'' を $W''_{3\alpha}$ の大きさが α となるように選べば，$W''_{3\alpha}$ はノンパラメトリック単純仮説 $\mathscr{H}(F_0;\mathscr{C}_1\cup\mathscr{C}'_1\cup F_0)$ に対する大きさ α の一致検定になる．すなわち $\mathscr{C}_1\cup\mathscr{C}'_1$ の任意の c.d.f. に対する $F_0(x)$ の大きさの α の一致検定である．

14.3　連続分布からの2標本問題

(a)　序　　論

$(x_1,\cdots,x_{n_1})$，$(x'_1,\cdots,x'_{n_2})$ をそれぞれ連続な c.d.f. $F_1(x), F_2(x)$ からの独立な標本とし，それぞれ O_{n_1}，O_{n_2} で表わす．ノンパラメトリック統計的推測の基本問題は O_{n_1}，O_{n_2} による情報に基づいて，$F_1(x) \doteqdot F_2(x)$ なる対 $(F_1(x), F_2(x))$ の種々のクラスにおける対立仮説に対する統計的仮説 $F_1(x) \equiv F_2(x)$ の一致検定方式を開発することである．

この種の検定方式は Dixon (1940)，Mathisen (1943)，Smirnov (1939 *a*)，Wald と Wolfowitz (1940)，Wilks (1961) 等により開発されている．本節ではこれらの検定のいくつかを考察しよう．

仮説 $F_1(x) \equiv F_2(x)$ の検定を論ずる場合，$F_1(x) \equiv F_2(x)$ という仮定に対して対立仮説のクラスを考えなければならない．便宜上，連続な c.d.f. の対 $(F_1(x), F_2(x))$ の全体を $\mathscr{J}$，$F_1(x)$ と $F_2(x)$ が一致する，すなわち $F_1(x) \equiv F_2(x)$ なる $\mathscr{J}$ の部分クラスを $\mathscr{J}_0$ で表わそう．$\mathscr{J}_0$ の要素における共通な c.d.f. を $F(x)$ とする．仮説 $F_1(x) \equiv F_2(x)$ に対する最も一般的な検定は，任意の要素 $(F_1(x), F_2(x)) \in \mathscr{J} - \mathscr{J}_0$ に対する $(F_1(x), F_2(x)) \in \mathscr{J}_0$ の検定であろう．14.3 (e) 節で考察した Smirnov (1939 *a*) 検定はこのような一般的な一致検定の例である．ここでは $\mathscr{J} - \mathscr{J}_0$ のある部分クラスの要素に対する仮説 $(F_1(x), F_2(x)) \in \mathscr{J}_0$ の一致検定だけを取り扱う．$\mathscr{J}^*$ をこのような部分クラスとすれば $(F_1(x), F_2(x)) \in \mathscr{J}^*$ に対する $(F_1(x), F_2(x)) \in \mathscr{J}_0$ の検定は $\mathscr{H}(\mathscr{J}^0;\mathscr{J}^0 \cup \mathscr{J}^*)$ で表わされる．

個々の検定を論ずる前に，同一で連続な c.d.f. からの2つの独立な標本に関する分布論を若干述べておくと有益である．

(b)　2標本の順序統計量によって生成された基本確率変数の分布論

同一で連続な c.d.f. を持つ母集団から2標本の順序統計量に関する単純ではあるが基本的な定理は次のように述べられる．

14.3.1　O_{n_1}，O_{n_2} はそれぞれ順序統計量として $(x_{(1)},\cdots,x_{(n_1)})$，$(x'_{(1)},\cdots,x'_{(n_2)})$ を持ち，同一で連続な c.d.f. からの独立な標本とし，$z_1,\cdots,z_{n_1+n_2}$ を2つの順序統計量を合わせて順序をつけた組とすれば

(14.3.1)
$$P(z_1 < \cdots < z_{n_1+n_2}) = 1\Big/\binom{n_1+n_2}{n_1}.$$

これを確かめよう．O_{n_1}，O_{n_2} が抽出された共通の c.d.f. を $F(x)$ とする．簡単のために2標本順序統計量の特定の“かみ合わせ”の確率，たとえば

(14.3.2)
$$P(x_{(1)} < \cdots < x_{(n_1)} < x'_{(1)} < \cdots < x'_{(n_2)})$$

を計算すれば十分である．(14.3.2) の確率変数を $z_1,\cdots,z_{n_1+n_2}$ に書き改め，$y_1 = F(z_1)$，$\cdots$，$y_{n_1+n_2} = F(z_{n_1+n_2})$ とすれば，(8.7.2) によって y の p.e. は

(14.3.3)
$$n_1!\,n_2!\,dy_1\cdots dy_{n_1+n_2}.$$

事象 $z_1 < \cdots < z_{n_1+n_2}$ が起るのは $y_1 < \cdots < y_{n_1+n_2}$ のとき，かつこのときに限る．この事象の確率は

(14.3.4)
$$n_1!\,n_2!\int_0^1\cdots\int_0^{y_3}\int_0^{y_2} dy_1\,dy_2\cdots dy_{n_1+n_2} = 1\Big/\binom{n_1+n_2}{n_1}.$$

(14.3.2) の2つの順序統計量の<u>任意</u>の並べ方に対しても (14.3.4) が成り立つことは明らかである．これで **14.3.1** を終わる．

前と同様に O_{n_1}，O_{n_2} をそれぞれ $F_1(x)$，$F_2(x)$ からの独立な標本とする．ただし $(F_1(x), F_2(x)) \in \mathscr{J}_0$ で $F(x)$ が共通の c.d.f. である．8.7 節で定義したように O_{n_1} の順序統計量は標本ブロック $B_1^{(1)},\cdots,B_1^{(n_1+1)}$ を定める．$F(x)$ で決まるこのブロックのカバーをそれぞれ $u_1,\cdots,u_{n_1+1}$ とする．

$r_{11},\cdots,r_{n_1+1}$ をそれぞれ $B_1^{(1)},\cdots,B_1^{(n_1+1)}$ にはいる O_{n_2} の要素の数とする．ただし $r_{n_1+1} = n_2 - r_1 - \cdots - r_{n_1}$ である．便宜上，確率変数 $(r_1,\cdots,r_{n_1+1})$ を O_{n_2} から決まる O_{n_1} の**ブロック頻度**と呼ぼう．このブロック頻度の分布を考察しよう．

さて $(u_1,\cdots,u_{n_1},r_1,\cdots,r_{n_1})$ は混合型 (n_1+n_2) 次元確率変数である．ただし u が連続型，r が離散型で，u の値を固定した r の条件につき p.f. は

(14.3.5)
$$\frac{n_2!}{r_1!\cdots r_{n_1+1}!}\,u_1^{r_1}\cdots u_{n_1+1}^{r_{n_1+1}}$$

である．しかし **8.7.4** より u の p.e. は

(14.3.6)
$$n_1!\,du_1\cdots du_{n_1}.$$

よって u と r の結合 p.f.-p.e. は (14.3.5) と (14.3.6) の積になる．r の（非条件つき）p.f. を得るには，この積を単体 $S_{n_1}: \{(u_1,\cdots,u_{n_1}): u_1 \geqslant 0,\cdots,u_{n_1} \geqslant 0,\ u_1+\cdots+u_{n_1} \leqslant 1\}$ 上で積分すればよい．7.7 節のディリクレ分布の性質を用いれば，この積分値は

n_1 次元確率変数 $(r_1,\cdots,r_{n_1})$ の p.f. として

(14.3.7)
$$1\Big/\binom{n_1+n_2}{n_1}$$

を与える．ただし各 r は $r_1+\cdots+r_{n_1}\leqq n_2$ となるすべて0か正整数である．

(14.3.7) を得た方法を同様に用いれば，確率変数 $(r_1,\cdots,r_{n_1})$ の任意 t 個（これを $(r_1^*,\cdots,r_t^*)$ で表わす）は p.f.

(14.3.8)
$$\binom{n_1+n_2-r_1^*-\cdots-r_t^*-t}{n_1-t}\Big/\binom{n_1+n_2}{n_1}$$

を持つことが示される．要約すれば，次の結果になる．

14.3.2　O_{n_1}, O_{n_2} を同一で連続な c.d.f. からの独立な標本とする．$(r_1,\cdots,r_{n_1+1})$ を O_{n_2} から決まる O_{n_1} のブロック頻度とする．このとき条件 $r_1+\cdots+r_{n_1+1}=n_2$ を満たす $(r_1,\cdots,r_{n_1+1})$ の p.f. は $(r_1,\cdots,r_{n_1+1})$ のすべての可能な標本点上で定数 $1\Big/\binom{n_1+n_2}{n_2}$ を持つ．さらにこの確率変数の任意 t 個，たとえば $(r_1^*,\cdots,r_t^*)$ の p.f. は (14.3.8) で与えられる．

$(r_1',\cdots,r_{n_2+1}')$ を O_{n_1} から決まる O_{n_2} のブロック頻度を表わす確率変数とすれば，$(r_1',\cdots,r_{n_2+1}')$ は p.f. としてこのすべての可能な標本点で定数

(14.3.9)
$$1\Big/\binom{n_1+n_2}{n_2}$$

を持つ．ただし $r_1'+\cdots+r_{n_2+1}'=n_1$ である．したがって，(14.3.7) と (14.3.9) は同じ値を持つので，$(r_1',\cdots,r_{n_2+1}')$ と $(r_1,\cdots,r_{n_1+1})$ の p.f. も2つの標本空間上で同じ定数を持つ．

(14.3.7) で与えられるブロック頻度数 $(r_1,\cdots,r_{n_1})$ の p.f. は (14.3.1) で与えられる O_{n_1} と O_{n_2} の順序統計量を合わせた異なる並べ方の p.f. と一致する．もちろんこれは驚くに値しない．なぜならベクトル確率変数 $(r_1,\cdots,r_{n_1+1})$ の値と O_{n_1}, O_{n_2} を合わせた順序統計量の並べ方との間に1対1の対応が存在するからである．

注意　O_{n_1}, O_{n_2} が抽出される連続な c.d.f. は k 次元でもよい．この場合，8.7 (c) 節で述べたように k 次元標本ブロックを切断する任意の方法が使えて，$B_k^{(1)},\cdots,B_k^{(n_1+1)}$ を n_2 個の点よりなる標本 O_{n_2} より定まるブロックとすることができる．このとき $(r_1,\cdots,r_{n_1+1})$ は R_k の n_2 個の点のうちそれぞれこのブロックに落ちる個数とすれば，確率変数 $(r_1,\cdots,r_{n_1+1})$ の分布は **14.3.2** と同じになる．

さて s_0 をブロック $B_1^{(1)},\cdots,B_1^{(n_1+1)}$ のうち O_{n_2} の要素を1つも含まない個数，s_1 を O_{n_2} の要素を1つ含むブロック数，以下同様に $s_2,\cdots,s_{n_2}$ を定義する．$s_0,s_1,\cdots,s_{n_2}$ は O_{n_2} より決まる O_{n_1} のブロック頻度数と呼ばれている．この $(s_0,s_1,\cdots,s_{n_2})$ は条件

(14.3.10)
$$\begin{aligned} s_0+s_1+\cdots+s_{n_2}&=n_1+1\\ s_1+2s_2+\cdots+n_2s_{n_2}&=n_2\end{aligned}$$

を満たす多次元確率変数である．確率変数 $(r_1,\cdots,r_{n_1+1})$ の標本空間の点はすべて等確率だから，$(s_0,\cdots,s_{n_2})$ の p.f. を求める問題は，一定の条件を満たす r の標本空間における点を列挙することにすぎない．簡単な考察より次の事柄がわかる．確率変数 $(s_0,s_1,\cdots,s_{n_2})$ の与えられた標本点 $(s_0',s_1',\cdots,s_{n_2}')$ に射影する確率変数 $(r_1,\cdots,r_{n_1+1})$ の標本点の数は

(14.3.11)
$$(u_0+u_1v+u_2v^2+\cdots+u_{n_2}v^{n_2})^{n_1+1}$$

の形式的展開における $u_0^{s_0'}u_1^{s_1'}\cdots u_{n_2}^{s_{n_2}'}v^{n_2}$ の係数である．この係数を $\binom{n_1+n_2}{n_1}$ で割って，ダッシュをのぞくと $(s_0,s_1,\cdots,s_{n_2})$ の p.f.

(14.3.12)
$$\frac{(n_1+1)!}{s_0!\,s_1!\cdots s_{n_2}!}\Big/\binom{n_1+n_2}{n_1}$$

を得る．ただし $s_0,s_1,\cdots,s_{n_2}$ は条件 (14.3.10) を満たす．(14.3.12) に含まれる s の個数は n_2 に依存することに注意しよう．指定した個数だけの，たとえば $(s_0,s_1,\cdots,s_t)$ の p.f. は (14.3.11) において $u_{t+1}=\cdots=u_{n_2}=u$ と置き $u_0^{s_0}u_1^{s_1}\cdots u_t^{s_t}u^sv^{n_2}$ の係数を選べば求められる．ただし $s_0+s_1+\cdots+s_t+s=n_1+1$, $s=s_{t+1}+\cdots+s_{n_2}$. $t=0$ のとき，この方法は s_0 の p.f.

(14.3.13)
$$p(s_0)=\frac{\binom{n_1+1}{s_0}\binom{n_2-1}{n_1-s_0}}{\binom{n_1+n_2}{n_1}}$$

を与える．s_0 の標本点は $k,k+1,\cdots,n_1$, $k=\max(0,n_1-n_2+1)$ である．t の値が大きいとき，p.f. は複雑になるので，省略しよう．

確率変数 $(s_0,s_1,\cdots,s_{n_2})$ の1つあるいは2つ以上の変数に対する階乗モーメントは (6.1.7) と同様にして，(14.3.12) から得られる．一般に

(14.3.14)
$$\mathscr{E}(s_0^{[g_0]}s_1^{[g_1]}\cdots s_t^{[g_t]})=(n_1+1)^{[g_0+\cdots+g_t]}\cdot\binom{n_1+n_2-g_0-2g_1-\cdots-(t+1)g_t}{n_1-g_0-\cdots-g_t}\Big/\binom{n_1+n_2}{n_1}$$

になる．ここでも，もちろん $g_0, \cdots, g_t$ は非負整数で，$g_0 + \cdots + g_t \leqslant n_1$, $g_1 + 2g_2 + \cdots + tg_t \leqslant n_2$ を満たす．(14.3.14) より s の1つあるいは2つ以上に関する平均，分散，共分散，その他のモーメント等が求められる．

要約すれば次の結果になる．

14.3.3　O_{n_1}, O_{n_2} が同一で連続な c.d.f. からの独立標本ならば，O_{n_2} より決まる O_{n_1} のブロック頻度数 $(s_0, s_1, \cdots, s_{n_2})$ の p.f. は (14.3.12) で与えられる．ただし $(s_0, s_1, \cdots, s_{n_2})$ の成分は (14.3.10) を満たす．さらに $(s_0, s_1, \cdots, s_t)$ の一般階乗モーメントは (14.3.14) で与えられる．

特に s_0 については

(14.3.15)
$$\mathscr{E}(s_0) = \frac{n_1(n_1+1)}{(n_1+n_2)}$$
$$\sigma^2(s_0) = \frac{n_1^2(n_1^2-1)}{(n_1+n_2)(n_1+n_2-1)} + \frac{n_1(n_1+1)}{(n_1+n_2)} - \frac{n_1^2(n_1+1)^2}{(n_1+n_2)^2}.$$

$n_2 = \rho n_1 + O(1)$, $\rho > 0$ とすれば，上式は

(14.3.16)
$$\mathscr{E}(s_0) = n_1\left(\frac{1}{1+\rho} + O\left(\frac{1}{n_1}\right)\right)$$
$$\sigma^2(s_0) = n_1\left(\frac{\rho^2}{(1+\rho)^3} + O\left(\frac{1}{n_1}\right)\right)$$

に帰着する．

さらに次の結果は大標本における s_0 の漸近分布の状態を述べている．

14.3.4　O_{n_1}, O_{n_2} は同一で連続な c.d.f. からの独立な標本で，n_1, n_2 が $n_2 = \rho n_1 + O(1)$ で十分に大きいとき，s_0 は漸近的に $N\left(\frac{n_1}{1+\rho}, \frac{n_1\rho^2}{(1+\rho)^3}\right)$ に従って分布する．

14.3.4 は (14.3.13) における $p(s_0)$ の階乗をすべてスターリングの公式で近似すれば証明される．これによると

(14.3.17)
$$s_0 = \frac{n_1}{1+\rho} + y\sqrt{\frac{n_1\rho^2}{(1+\rho)^3}}.$$

よって

(14.3.18)
$$\lim_{n_1\to\infty} \sum_{s_0=s_0'(n_1)}^{s_0''(n_1)} p(s_0) = \frac{1}{\sqrt{2\pi}}\int_{y'}^{y''} e^{-\frac{1}{2}t^2}\,dt.$$

ただし $s_0'(n_1)$, $s_0''(n_1)$ は (14.3.17) において右辺の y にそれぞれ y', y'' を代入して得られる数以下の最大整数である．以下詳細は省略する．

(c)　空ブロック検定

まずノンパラメトリック2標本検定として簡単な**空ブロック検定**を考察しよう．この検定は14.2(c)節で論じられたデビッドの1標本空細胞検定に影響されている．検定は棄却域 $W_{a\alpha}$ として，$s_0 \geqslant s_0(\alpha, n_1, n_2)$ を満たす s_0 の値の集合を持つ．ただし $s_0(\alpha, n_1, n_2)$ は

(14.3.19)
$$P(W_{a\alpha} \mid (F_1, F_2) \in \mathscr{J}_0) \leqslant \alpha$$

となる最小整数である．

$(F_1, F_2) \in \mathscr{J}_0$ すなわち $F_1(x) \equiv F_2(x)$ のとき，s_0 の p.f. は (14.3.13) で与えられる．よって小さい n_1, n_2 に対する $s_0(\alpha, n_1, n_2)$ の値はこの p.f. から求められる．

n_1, n_2 が $n_2 = \rho n_1 + O(1)$ を満たして，大きいとき **14.3.4** から棄却値 $s_0(\alpha, n_1, n_2)$ を次のように近似できる．

(14.3.20)
$$s_0(\alpha, n_1, n_2) = \frac{n_1}{1+\rho} + y_\alpha\sqrt{\frac{n_1\rho^2}{(1+\rho)^3}} + O(1).$$

ただし

(14.3.21)
$$\frac{1}{\sqrt{2\pi}}\int_{y_\alpha}^{\infty} e^{-\frac{1}{2}t^2}\,dt = \alpha.$$

よって **14.3.4** より，$n_2/n_1 \longrightarrow \rho > 0$ で $n_1, n_2 \longrightarrow \infty$ ならば

(14.3.22)
$$\lim P(W_{a\alpha} \mid (F_1, F_2) \in \mathscr{J}_0) = \alpha$$

が成り立つ．(F_1, F_2) が $\mathscr{J}_0$ に属さないとき，s_0 の分布は s_0 軸の右側にかたよる傾向があることからすれば，この検定 $W_{a\alpha}$ は妥当である．

この妥当性をより形式的に調べるために，$W_{a\alpha}$ の一致性を調べよう．$F_1^{-1}(u)$ を c.d.f. $F_1(x)$ の逆関数，$\mathscr{J}^*$ を次の (i), (ii) を満たす連続な c.d.f. の対 $(F_1(x), F_2(x))$ のクラスとする．

（i）$F_2(F_1^{-1}(u))$ は確率0の集合をのぞく $(0,1)$ 上のすべての u について導関数 $g(u)$ を持つ．

（ii）$(0,1)$ 上での u に関する $F_2(F_1^{-1}(u))$, $F_1(F_1^{-1}(u))\ (\equiv u)$ の導関数は確率正の集合上で異なる．

次の結果を証明しよう．

14.3.5　検定 $W_{a\alpha}$ は $n_2=\rho n_1+O(1)$ $(\rho>0)$ で，$n_1, n_2\longrightarrow\infty$ のとき任意の $(F_1, F_2)\in\mathscr{J}^*$ に対する任意の $(F_1, F_2)\in\mathscr{J}_0$ の一致検定である．

14.3.5 の証明には，$(F_1, F_2)\in\mathscr{J}^*$ ならば，$n_2/n_1\longrightarrow\rho>0$ となるように $n_1, n_2\longrightarrow\infty$ のとき，$s_0/(n_1+1)$ が $1/(1+\rho)$ より大きい数に確率収束することを示せばよい．$1/(1+\rho)$ は $(F_1, F_2)\in\mathscr{J}_0$ のとき，$s_0/(n_1+1)$ が確率収束する値である．

z_ξ $(\xi=1, \cdots, n_1+1)$ を O_{n_1} より定まる標本ブロック $B_1^{(\xi)}$ が O_{n_2} の要素を1つも含まないとき1，その他では0をとる確率変数とする．このとき

(14.3.23)
$$s_0=z_1+\cdots+z_{n_1+1}.$$

O_{n_1}，O_{n_2} はそれぞれ $\mathscr{J}^*$ に属す c.d.f. の対 (F_1, F_2) からの独立な標本であることを仮定して $\mathscr{E}(s_0)$ を計算しよう．

$$\mathscr{E}(z_\xi)=P(z_\xi=1)$$

より，$\mathscr{E}[s_0/(n_1+1)|(F_1, F_2)\in\mathscr{J}^*]$ を $\mathscr{E}[s_0/(n_1+1)|\mathscr{J}^*]$ で表わせば

(14.3.24)
$$\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right)=\frac{1}{n_1+1}[P(z_1=1)+\cdots+P(z_{n_1+1}=1)].$$

ところが $\xi=2, \cdots, n_1$ について

(14.3.25)
$$P(z_\xi=1)$$
$$=\frac{n_1!}{(\xi-2)!(n_1-\xi)!}\int_{S_2}[1-F_2(x_{(\xi)})+F_2(x_{(\xi-1)})]^{n_2}[F_1(x_{(\xi-1)})]^{\xi-2}$$
$$\cdot[1-F_1(x_{(\xi)})]^{n_1-\xi}dF_1(x_{(\xi-1)})dF_1(x_{(\xi)})$$

である．ただし S_2 は $-\infty<x_{(\xi-1)}<x_{(\xi)}<+\infty$ なるユークリッド空間 R_2 の領域．

$P(z_1=1)$ および $P(z_{n_1+1}=1)$ については，(14.3.25) よりもやや簡単に表わされる．しかし n_1+1 で割って $n_2/n_1\longrightarrow\rho>0$，$n_1, n_2\longrightarrow\infty$ とすれば，ともに極限は0である．

次の変数変換を行なおう．

$$u_1=F_1(x_{(\xi-1)}),\qquad u_2=F_1(x_{(\xi)}).$$

このとき (14.3.25) は

(14.3.26)
$$P(z_\xi=1)$$
$$=\frac{n_1!}{(\xi-2)!(n_1-\xi)!}\int_{T_2}[1-G(u_1)+G(u_2)]^{n_2}u_1^{\xi-2}(1-u_2)^{n_1-\xi}du_1du_2$$

になる．ただし T_2 は3角形 $0<u_1<u_2<1$ で，$G(u)=F_2(F_1^{-1}(u))$．$G(u)$ は導関数 $g(u)$ を持つと仮定している．このとき $g(u)$ は $(0,1)$ 上の確率密度関数である．

よって

(14.3.27)
$$\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right)=\frac{1}{n_1+1}\sum_{\xi=2}^{n_1}P(z_\xi=1)+\delta_{n_1,n_2}.$$

ただし $\delta_{n_1,n_2}=(n_1+1)^{-1}[P(z_1=1)+P(z_{n_1+1}=1)]$．(14.3.26) の $P(z_\xi=1)$ を (14.3.27) に代入して和をとれば

(14.3.28)
$$\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right)=\frac{n_1(n_1-1)}{n_1+1}\int_{T_2}[1-u_2+u_1]^{n_1-2}$$
$$\cdot[1-G(u_2)+G(u_1)]^{n_2}du_1du_2+\delta_{n_1,n_2}.$$

さて，変換
$$u_1=v,\qquad u_2=u+\frac{y}{n_1}$$
を行なうと (14.3.28) は

(14.3.29)
$$\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right)=\frac{n_1(n_1-1)}{n_1(n_1+1)}\int_0^1\int_0^{(1-v)n_1}\left[1-\frac{y}{n_1}\right]^{n_1-2}$$
$$\cdot\left[1-\left\{\frac{G\left(v+\dfrac{y}{n_1}\right)-G(v)}{y/n_1}\right\}\frac{y}{n_1}\right]dy\,dv+\delta_{n_1,n_2}.$$

$n_2/n_1\longrightarrow\rho>0$ となるように $n_1, n_2\longrightarrow\infty$ として極限をとれば

(14.3.30)
$$\lim\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right)=\int_0^1\int_0^\infty e^{-y[1+\rho g(v)]}dy\,dv.$$

y で積分して

(14.3.31)
$$\lim\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right)=\int_0^1\frac{dv}{1+\rho g(v)}$$

を得る．ここで $g(v)$ が確率正の集合上で1と異なるとき，(14.3.31) の右辺の積分が $1/(1+\rho)$ を超えることを示さねばならない．すなわちこれは $(F_1, F_2)\in\mathscr{J}^*$ のとき $F_2(F_1^{-1}(u))$ と u の導関数は $(0,1)$ 上の確率正の集合上で異なるからである．

この条件の下では次のように厳密な意味でシュワルツの不等式が成立する．

(14.3.32)
$$\int_0^1\frac{dv}{1+\rho g(v)}\cdot\int_0^1[1+\rho g(v)]dv>\left[\int_0^1\sqrt{\frac{1}{1+\rho g(v)}}\sqrt{1+\rho g(v)}\,dv\right]^2.$$

しかし不等式の右辺は1，左辺の第2項は値 $1+\rho$ を持つ．したがって

(14.3.33)
$$\int_0^1\frac{dv}{1+\rho g(v)}>\frac{1}{1+\rho}.$$

すなわち $n_2=\rho n_1+O(1)$，$\rho>0$ で $n_1, n_2\longrightarrow\infty$ とすれば

$$\lim\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right)>\lim\mathscr{E}\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}_0\right)$$

が成り立つ．

空細胞検定の一致性の証明を終えるには，$n_2 = n_1\rho + O(1)$ なる大きな n_1，n_2 について，$(F_1, F_2) \in \mathscr{J}^*$ に関する $s_0/(n_1+1)$ の分散の次数が $1/n_1$ であることを示せばよい．$\|\sigma_{\xi\eta}\|$ を $(z_1, \cdots, z_{n_1+1})$ の共分散行列とすれば

(14.3.34) $$\sigma^2\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right) = \frac{1}{(n_1+1)^2}\left[\sum_{\xi,\eta=1}^{n_1+1}\sigma_{\xi\eta}\right].$$

ただし z_ξ と z_η の共分散 $\sigma_{\xi\eta}$ は

(14.3.35) $$\sigma_{\xi\eta} = P(z_\xi = 1, z_\eta = 1) - P(z_\xi = 1)P(z_\eta = 1)$$

で表わされる．すべての $\xi \neq \eta$ に対して $\sigma_{\xi\eta} < 0$ を示そう．最初の標本 O_{n_1} で生成される基本カバー $u_1, \cdots, u_{n_1}$ は (14.3.6) の確率素分 $n_1!\,du_1 \cdots du_{n_1}$ を持つ．この微分を $dH(u)$ で表わす．

$p_\xi = F_2(x_{(\xi)}) - F_2(x_{(\xi-1)})$ とし，p_η も同様に定義する．$u_1 + \cdots + u_\xi = F_1(x_{(\xi)})$，$\xi = 1, \cdots, n_1$ は $x_{(\xi)}$ と u_ξ との1対1変換だから，確率変数 $p_1, \cdots, p_{n_1}$ は明らかに $u_1, \cdots, u_{n_1}$ の関数である．与えられた最初の標本に対して，次の条件つき確率が得られる．

(14.3.36) $$P(z_\xi = 1, z_\eta = 1 | u_1, \cdots, u_{n_1}) = (1 - p_\xi - p_\eta)^{n_2}.$$

非条件つき確率は

(14.3.37) $$P(z_\xi = 1, z_\eta = 1) = \mathscr{E}(1 - p_\xi - p_\eta)^{n_2}.$$

ただし $\mathscr{E}$ は $(u_1, \cdots, u_{n_1})$ の分布に関してとる．

同様に

(14.3.38) $$\begin{aligned} P(z_\xi = 1 | u_1, \cdots, u_{n_1}) &= (1 - p_\xi)^{n_2} \\ P(z_\eta = 1 | u_1, \cdots, u_{n_1}) &= (1 - p_\eta)^{n_2}. \end{aligned}$$

非条件つき確率 $P(z_\xi = 1)$ および $P(z_\eta = 1)$ については

(14.3.39) $$P(z_\xi = 1)P(z_\eta = 1) = \mathscr{E}(1 - p_\xi)^{n_2} \cdot \mathscr{E}(1 - p_\eta)^{n_2}.$$

$P(z_\xi = 1)$ に対する式 (14.3.26) と同様な方法を用いれば

$$P(z_\xi = 1, z_\eta = 1), \qquad \xi \neq \eta.$$

すなわち

$$\mathscr{E}[(1 - z_\xi)^{n_2}(1 - z_\eta)^{n_2}]$$

を4次元積分で表わすことができる．相当な計算（ここでは省略する．Blum と Weiss (1957) を参照.）によって

(14.3.40) $$\mathscr{E}(1 - p_\xi - p_\eta)^{n_2} \leqslant \mathscr{E}[(1 - p_\xi)^{n_2}] \cdot \mathscr{E}[(1 - p_\eta)^{n_2}]$$

が示された．しかし (14.3.40) は

(14.3.41) $$P(z_\xi = 1, z_\eta = 1) - P(z_\xi = 1) \cdot P(z_\eta = 1) \leqslant 0$$

と同値である．

したがって (14.3.35) は

(14.3.42) $$\sigma_{\xi\eta} \leqslant 0, \qquad \xi \neq \eta = 1, \cdots, n_1.$$

よって (14.3.34) は

(14.3.43) $$\sigma^2\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right) \leqslant \frac{1}{(n_1+1)^2}\sum_{\xi=1}^{n_1+1}(\sigma_{\xi\xi})$$

になる．ところが $\xi = 1, \cdots, n_1 + 1$ について

(14.3.44) $$\sigma_{\xi\xi} = P(z_\xi = 1) - [P(z_\xi = 1)]^2 = P(z_\xi = 1)[1 - P(z_\xi = 1)] \leqslant \frac{1}{4}$$

を用いれば

(14.3.45) $$\sigma^2\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right) \leqslant \frac{1}{4(n_1+1)}$$

(14.3.46) $$\sum_{i=0}^{k}\frac{(s_i - na_i)^2}{na_i} + \frac{u^2 + v^2}{n\rho^2(1+\rho)a_k} > \chi^2_\alpha$$

を得る．ただし $a_i = \rho^i/(1+\rho)^{i+1}$ で

(14.3.47) $$\begin{aligned} u &= \sum_{i=0}^{k}(s_i - na_i)(i - \rho - k - 1) \\ v &= \sqrt{\rho(1+\rho)}\sum_{i=0}^{k}(s_i - na_i). \end{aligned}$$

ゆえに $n_1 \longrightarrow \infty$ のとき $\sigma^2\left(\frac{s_0}{n_1+1}\middle|\mathscr{J}^*\right) \longrightarrow 0$．これで **14.3.5** の議論を終わる．

Blum と Weiss (1957) は任意のブロック頻度数 s_i が仮説 $(F_1, F_2) \in \mathscr{J}_0$ の一致検定になるような $\mathscr{J}^*$ の部分クラスを考えた．Wilks (1961) は k を固定して大標本について $s_0, s_1, \cdots, s_k$ に基づくカイ2乗に類似した検定 $W^*_{\alpha\alpha}$ を考察している．この $W^*_{\alpha\alpha}$ は次を満たす2標本空間内の棄却域を持つ．すなわち

$$\sum_{i=0}^{k}\frac{(s_i - na_i)^2}{na_i} + \frac{u^2 + v^2}{n\rho^2(1+\rho)a_k} > \chi^2_\alpha.$$

ただし $a_i = \rho^i/(1+\rho)^{i+1}$ で

$$u = \sum_{i=0}^{k}(s_i - na_i)(i - \rho - k - 1)$$

$$v = \sqrt{\rho(1+\rho)} \sum_{i=0}^{k} (s_i - na_i).$$

χ^2_α はカイ2乗分布 $C(k+1)$ の $100\alpha\%$ 点．この検定は当然のように s_0 だけに基づく検定 $W_{a\alpha}$ よりもいくぶん強力である．

注意　検定 $W_{a\alpha}$（あるいは $W^*_{a\alpha}$）は連続な c.d.f. を持つ k 次元分布の場合にまで拡張できる．このとき O_{n_1} は8.7 (c) 節で述べたように k 次元標本ブロック $B_k^{(1)}, \cdots, B_k^{(n_1+1)}$ を生成する．このブロックは用いられる順序関数系に依存する特定の順序に従って生成される．一方 O_{n_2} はブロックにはいる点を生成する．ブロック列にはいる点の個数は確率変数 $(r_1, \cdots, r_{n_1+1})$ の成分になる．ただし $r_1 + \cdots + r_{n_1+1} = n_2$．空ブロックの個数が確率変数 s_0 で，この分布は標本の c.d.f. が同一であるとき (14.3.13) で与えられる．k 次元の場合も c.d.f. に関するある特定の条件の下では1次元と同様にして一致検定であることが示される．

(d) 連 検 定

O_{n_1} の各順序統計量を C，O_{n_2} の各順序統計量を $\overline{C}$ で表わせば，O_{n_1} の順序統計量と O_{n_2} の順序統計量をすべて合わせた統計量の任意の配列は n_1 個の C と n_2 個の $\overline{C}$ の置換に対応する．さて **14.3.1** の条件のとき，n_1 個の C と n_2 個の $\overline{C}$ の $\binom{n_1+n_2}{n_1}$ 通りの可能な配列はすべて等確率である．よって6.6節における連の理論がすべて，C，$\overline{C}$ による $\binom{n_1+n_2}{n_1}$ 通りの可能な列によって生成される連に適用できる．（6.6節では $n_1 + n_2$ を n で表わしている．）

2標本検定として Wald と Wolfowitz (1940) が提唱した統計的関数は6.6 (a) 節で定義した C と $\overline{C}$ の列における（すなわち，O_{n_1} と O_{n_2} のそれぞれの順序統計量を結合した列における）連の総数 u である．$(F_1(x), F_2(x)) \in \mathscr{J}_0$ のとき，u の p.f. は (6.6.20) で与えられる．一方 $\mathscr{E}(u)$ と $\sigma^2(u)$ は (6.6.19) で与えられる．確率変数 u と s_0 は強い負の相関があることに注意されたい．

u の値が小さいとき，仮説 $(F_1(x), F_2(x)) \in \mathscr{J}_0$ の真実性が疑われる傾向にあるので u の棄却域として

(14.3.48)
$$P(u \leqslant u(\alpha, n_1, n_2)) \leqslant \alpha$$

を満たす集合を定義する．ただし $u(\alpha, n_1, n_2)$ は (14.3.48) が成立する最大整数である．この検定を $W_{b\alpha}$ と呼ぼう．検定の大きさは近似的に α である．

n_1 と n_2 の値が大きいとき，u は漸近的に正規分布に従い，平均と分散が (6.6.19)

で与えられる．ワルドとウォルフォヴィッツが次の定理で明確に述べている．

14.3.6　$(F_1(x), F_2(x)) \in \mathscr{J}_0$ のとき，$n_2/n_1 \longrightarrow \rho > 0$ となるように $n_1, n_2 \longrightarrow \infty$ として極限をとれば，u は漸近的に

$$N\left(\frac{2n_1\rho}{1+\rho}, \frac{4\rho^2 n_1}{(1+\rho)^3}\right)$$

に従う．

14.3.6 の証明は **9.6.3** と同様である．ここでは簡単に概要を述べるだけで，詳細は省こう．

$$y = \frac{\left(u - \dfrac{2n_1\rho}{1+\rho}\right)(1+\rho)^{\frac{3}{2}}}{2\rho\sqrt{n_1}}$$

とすれば

(14.3.49)
$$u \doteq \frac{2\rho\sqrt{n_1}(y + \sqrt{n_1}(1+\rho))}{\sqrt{(1+\rho)^3}}$$

である．さて (6.6.20) の $p(u)$ を，偶数のときは $p_1(u)$ で，奇数のときは $p_2(u)$ でそれぞれ表わす．n_1 と $t' < t''$ を任意に与えれば

(14.3.50)
$$P(y' < y < y'') = \sum\nolimits_1 p_1\left(\frac{2\rho\sqrt{n_1}(y + \sqrt{n_1}(1+\rho))}{\sqrt{(1+\rho)^3}}\right) + \sum\nolimits_2 p_2\left(\frac{2\rho\sqrt{n_1}(y + \sqrt{n_1}(1+\rho))}{\sqrt{(1+\rho)^3}}\right).$$

ただし $\sum_1$ は，u が偶数のとき (14.3.48) で与えられる y' と y'' の間の y の離散値にわたる和，$\sum_2$ は，u が奇数のとき (14.3.48) で与えられる同様な和である．

さて最初の和を見よう．これは

(14.3.51)
$$\sum\nolimits_1 \left[p_1\left(\frac{2\rho n_1(y + \sqrt{n_1}(1+\rho))}{\sqrt{(1+\rho)^3}}\right)\sqrt{n_1\rho^2/(1+\rho)^3}\right]\Delta y$$

と書かれる．ただし $\Delta y = \sqrt{(1+\rho)^3/(n_1\rho^2)}$．$p_1(u)$（すなわち u が偶数のときの $p(u)$）に含まれるすべての階乗に対して，大きな階乗に対するスターリングの公式 (7.6.29) を適用すれば，$n_1 \longrightarrow \infty$ のとき (14.3.51) は

(14.3.52)
$$\frac{1}{2\sqrt{2\pi}} \int_{y'}^{y''} e^{-\frac{1}{2}t^2}\, dt$$

に収束することがわかる．同様に (14.3.50) の2番目の項も，$n_1 \longrightarrow \infty$ のとき (14.3.52) の積分に収束する．よってこの2つの極限を合わせると

(14.3.53)
$$\lim_{n_1\to\infty} P(y'<y<y'')=\frac{1}{\sqrt{2\pi}}\int_{y'}^{y''} e^{-\frac{1}{2}t^2}\,dt$$

が任意の固定された $y'<y''$ に対して成り立つ．これで **14.3.6** の証明を終わる．

$\lim n_2/n_1=\rho>0$ になるように $n_1, n_2\longrightarrow\infty$ とし，$(F_1(x), F_2(x))\in\mathscr{J}_0$ ならば **14.3.6** より

(14.3.54)
$$\lim_{n_1\to\infty} P(v\leqslant u(\alpha, n_1, \rho n_1))=\alpha$$

が成り立つ．

Wald と Wolfowitz (1940) は $W_{a\alpha}$ 検定が **14.3.5** におけるクラス $\mathscr{J}^*$ の要素に対する $(F_1(x), F_2(x))\in\mathscr{J}_0$ の一致検定であることを示している．

(e) スミルノフ検定

$F_{1n_1}(x)$，$F_{2n_2}(x)$ は (11.6.2) で定義されたように，それぞれ O_{n_1}，O_{n_2} で定まる経験 c.d.f. とする．Smirnov (1939 *a*) 統計量

(14.3.55)
$$D_{n_1,n_2}=\sup_x |F_{1n_1}(x)-F_{2n_2}(x)|$$

はそれ自身仮説 $(F_1(x), F_2(x))\in\mathscr{J}_0$ の検定に対する妥当な基準であることを示している．

もっと明確に述べよう．$W_{c\alpha}$ を O_{n_1} と O_{n_2} の標本空間における

(14.3.56)
$$P(D_{n_1,n_2}>\delta(\alpha, n_1, n_2))\leqslant\alpha$$

なる棄却域とする．ただし $\delta(\alpha, n_1, n_2)$ は $(F_1(x), F_2(x))\in\mathscr{J}_0$ のとき (14.3.56) が成立する最小数として選ばれる．

スミルノフは $(F_1(x), F_2(x))\in\mathscr{J}_0$，$n_2/n_1\longrightarrow\rho>0$，$n_1, n_2\longrightarrow\infty$ ならば

(14.3.57)
$$\lim_{n_1,n_2\to\infty} P(D_{n_1,n_2}>\lambda\sqrt{1/n_1+1/n_2})=1-\sum_{i=-\infty}^{\infty}(-1)^i e^{-2i^2\lambda^2}$$

であることを示した．よって λ_α を

(14.3.58)
$$\sum_{i=-\infty}^{\infty}(-1)^i e^{-2i^2\lambda_\alpha^2}=1-\alpha$$

を満たすとして，$\delta(\alpha, n_1, n_2)=\lambda_\alpha\sqrt{1/n_1+1/n_2}$ にとれば，$n_1, n_2\longrightarrow\infty$，$n_2/n_1\longrightarrow\rho>0$ のとき

(14.3.59)
$$\lim_{n_1,n_2\to\infty} P\left(D_{n_1,n_2}>\lambda_\alpha\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=\alpha$$

が成り立つ．

さて $W_{c\alpha}$ は $(F_1(x), F_2(x))\in\mathscr{J}-\mathscr{J}_0$ に対する $(F_1(x), F_2(x))\in\mathscr{J}_0$ の一致検定，すなわち $F_1(x)\neq F_2(x)$ ならば，$n_1, n_2\longrightarrow\infty$，$n_2/n_1\longrightarrow\rho>0$ のとき

(14.3.60)
$$\lim_{n_1,n_2\to\infty} P\left(D_{n_1,n_2}>\lambda_\alpha\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)=1$$

であることが示される．この証明は省略する．

確率変数 D_{n_1,n_2} はブロック頻度 $(r_1, \cdots, r_{n_1+1})$ の関数である．したがって小標本のとき，任意の d に対して $P(D_{n_1,n_2}\leqslant d)$ は $D_{n_1,n_2}\leqslant d$ となる $(r_1, \cdots, r_{n_1+1})$ の標本空間内の標本点を数え上げて，$\binom{n_1+n_2}{n_1}$ で割って求められる．これができるのは，むしろ $n_1\fallingdotseq n_2$ のときである．しかし $n_1=n_2=n$ ならば $D_{n,n}$ は $1/n$ の倍数である．この場合 Massey (1951) は **11.7.1** と同様にして

(14.3.61)
$$P\left(D_{n_1,n}\leqslant\frac{k}{n}\right)=\sum_{i=0}^{k}U(i, n)\bigg/\binom{2n}{n}$$

を示した．ただし $U(\xi, \eta)$，$\xi=0, 1, \cdots, 2k-1$，$\eta=1, \cdots, n$ は $F_{1n}(x_{(\eta)})=(\xi+\eta-k)/n$ かつ $x<x_{(\eta)}$ のとき $|F_{1n}(x)-F_{2n}(x)|\leqslant k/n$ を満たす $(r_1, \cdots, r_{n_1+1})$ の標本空間内の標本点の個数である．($x_{(\eta)}$ は O_{n_1} の η 番目の順序統計量.) $U(\xi, \eta)$ は差分方程式

(14.3.62)
$$\begin{aligned}
U(0, \eta+1)&=U(0, \eta)+U(1, \eta)\\
U(1, \eta+1)&=U(0, \eta)+U(1, \eta)+U(2, \eta)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
U(2k-2, \eta+1)&=U(0, \eta)+\cdots+U(2k-1, \eta)\\
U(2k-1, \eta+1)&=U(0, \eta)+\cdots+U(2k-1, \eta)
\end{aligned}$$

を満たす．この境界条件は

(14.3.63)
$$\begin{aligned}
U(i, 0)&=0, \qquad i=1, \cdots, k-1\\
U(k, 0)&=1.
\end{aligned}$$

$n_1\neq n_2$ のときも同様な式が展開される．このとき D_{n_1,n_2} は $1/n_1n_2$ の倍数である．Massey (1951) はこのような再起式を用いて

$$P\left(D_{n_1,n_2}\leqslant\frac{k}{n_1n_2}\right), \qquad k=1, 2, \cdots, n_1n_2, \qquad n_1\leqslant n_2\leqslant 10$$

の表を作成している．

$n_1=n_2=n$ についての $P(D_{n_1,n_2}\leqslant k/n_1n_2)$ に対する比較的簡単な公式が Gnedenko と Koroliuk (1951) により見い出されている．これはここで述べるに値するだろう．

Gnedenko と Koroliuk (1951) の結果を得るには，まず確率変数 D_n^+，D_n^- を

(14.3.64)
$$D_n^+=\sup_x (F_{1n}(x)-F_{2n}(x))$$

(14.3.64 a)　$$D_n^- = \inf_x (F_{1n}(x) - F_{2n}(x))$$

で定義しよう．2標本を合わせた順序統計量を $z_{(1)} < z_{(2)} < \cdots < z_{(2n)}$ とし，ζ_t ($t=1, \cdots, 2n$) を $z_{(t)}$ が x（すなわち最初の標本に属す）ならば $\frac{1}{n}$，z_t が x'（すなわち次の標本に属す）ならば $-\frac{1}{n}$ をとる確率変数とする．

(14.3.65)　$$s_t = \zeta_1 + \cdots + \zeta_t$$

とすれば

(14.3.66)　$$D_n^+ = \sup_t s_t, \qquad D_n^- = \inf_t s_t$$

である．$s_0 = 0$ とし，(t, s) 平面において点列 (t, s_t), $t = 0, 1, \cdots, 2n$ からなるグラフを考えれば，$(0,0)$ を始点，$(2n, 0)$ を終点とするパスを得る．t の任意の連続した2つの整数値間では，パスは上がるか下がるかである．$\binom{2n}{n}$ 個の可能なパスがあり，1つのパスは2標本の順序統計量 $z_{(1)} < \cdots < z_{(2n)}$ における x, x' による可能な列に対応している．条件 $(F_1(x), F_2(x)) \in \mathscr{J}_0$ の下では，これらのパスはすべて等確率である．したがって，$P(D_n^+ \leqslant k/n)$ の値を求める問題は，直線 $s = (k+1)/n$ の下にすべて含まれてしまうパスの数を求めて，これを $\binom{2n}{n}$ で割ればよい．同様に $P(D_{n,n} \leqslant k/n)$ の値を求める問題は2直線 $s = \pm(k+1)/n$ の間にすべて含まれるパスの数を求めて $\binom{2n}{n}$ で割ればよい．この直線をそれぞれ L^+, L^- で表わそう．まず $P(D_n^+ \leqslant k/n)$ を考える．直線 L^+ の下に含まれてしまわない任意のパスを調べる．直線 L^+ と最初に交わる点 A が存在する．さてパスの A から $(2n, 0)$ までの部分に対応してこのパスの L^+ に関する鏡像は A から $(2n, 2(k+1)/n)$ を通る．よって $(0,0)$ を通り，直線 L^+ と交わり，$(2n, 0)$ で終わるパスの数は $(0,0)$ から始まり，L^+ を通過して，$(2n, 2(k+1)/n)$ を通る像パスの数に等しい．このパスの全体は，$n+k+1$ 回上向きになり，$n-k-1$ 回下向きになるようなパスの全体，すなわち $\binom{2n}{n+k+1}$ 個ある．よって，直線 L^+ と交わらないパスの個数は $\binom{2n}{n} - \binom{2n}{n+k+1}$．ゆえに

(14.3.67)　$$P\left(D_n^+ \leqslant \frac{k}{n}\right) = 1 - \frac{\binom{2n}{n+k+1}}{\binom{2n}{n}}.$$

対称性より

(14.3.68 a)　$$P\left(D_n^- \geqslant -\frac{k}{n}\right) = P\left(D_n^+ \leqslant \frac{k}{n}\right).$$

$\lambda = k/\sqrt{2n}$ とすれば

(14.3.69)　$$P(\sqrt{n/2}\, D_n^+ \leqslant \lambda) = 1 - \frac{\binom{2n}{n+\lambda\sqrt{2n}+1}}{\binom{2n}{n}}.$$

大きな階乗に対するスターリングの公式 (7.6.29) を用いれば

(14.3.70)　$$\lim_{n\to\infty} \frac{\binom{2n}{n+\lambda\sqrt{2n}+1}}{\binom{2n}{n}} = e^{-2\lambda^2}.$$

よって

(14.3.71)　$$\lim_{n\to\infty} P(\sqrt{n/2}\, D_n^+ < \lambda) = 1 - e^{-2\lambda^2}$$

である．要約すると次のようになる．

14.3.7　$F_{1n}(x)$, $F_{2n}(x)$ を同一で連続な c.d.f. からのおのおの大きさ n の2つの独立な標本における経験 c.d.f. とし，D_n^+, D_n^- を (14.3.64), (14.3.64 a) で定義する．このとき $P(D_n^+ \leqslant k/n)$ および $P(D_n^- \geqslant -k/n)$ は同じ値 $P_n^*(k/\sqrt{2}n)$ を持つ．この値は

(14.3.72)　$$P_n^*\left(\frac{k}{\sqrt{2}n}\right) = 1 - \frac{\binom{2n}{n+k+1}}{\binom{2n}{n}}.$$

しかも

(14.3.73)　$$\lim_{n\to\infty} P_n^*\left(\frac{\lambda}{\sqrt{n}}\right) = 1 - e^{-2\lambda^2}.$$

読者は $P_n(\lambda/\sqrt{n})$（ただし $P_n(d)$ は1標本に対しては (11.6.5) で定義されている）は $n \longrightarrow \infty$ のとき $P_n^*(\lambda/\sqrt{n})$ の極限と同じ値を持つことに注意されたい．

さて $P(D_{n,n} \leqslant k/n)$ の値を求めよう．S_0^+ を L^+ と交わるパスの集合，S_0^- を L^- と交わるパスの集合とすれば，$S_0^+ \cup S_0^-$ は L^+ と L^- の間に含まれてしまわないパス全体からなる．$N(E)$ をパスの任意の集合 E に含まれるパスの個数とする．このとき

(14.3.74)　$$N(S_0^+ \cup S_0^-) = N(S_0^+) + N(S_0^-) - N(S_0^+ \cap S_0^-).$$

ただし

(14.3.75)　$$N(S_0^+) = N(S_0^-) = \binom{2n}{n+k+1}.$$

さて $S_0^+ \cap S_0^-$ は2つの集合 S_1^+, S_1^- の和集合からなる．ただし S_1^+ は少なくとも L^+ を L^- に結合させている部分を含むパス全体からなり，S_1^- も同様に少なくとも L^- を L^+ に結合させている部分を含むパス全体からなっている．よって

(14.3.76)　$N(S_0^+ \cap S_0^-) = N(S_1^+ \cup S_1^-) = N(S_1^+) + N(S_1^-) - N(S_1^+ \cap S_1^-)$.

$N(S_0^+)$, $N(S_0^-)$ を求めたようにすれば

(14.3.77)
$$N(S_1^+) = N(S_1^-) = \binom{2n}{n+2(k+1)}.$$

S_i^+ を，L^+ を L^- に結合している部分を少なくとも i 個（いずれも L^+ から始まる）含むパスの集合として定義する．同様に S_i^- を定義すれば

(14.3.78)　$N(S_0^+ \cup S_0^-)$
$$= 2\left[\binom{2n}{n+k+1} - \binom{2n}{n+2(k+1)} + \cdots + (-1)^{r-1}\binom{2n}{n+r(k+1)}\right]$$

ただし $r = [n/(k+1)]$ は $n/(k+1)$ 以下の最大整数を表わす．

$N(D_{n,n} \leqslant k/n) = \binom{2n}{n} - N(S_0^+ \cup S_0^-)$ だから $\binom{2n}{n}$ で割ると

(14.3.79)　$P\left(D_{n,n} \leqslant \frac{k}{n}\right)$
$$= 1 + 2\sum_{i=1}^{r}(-1)^i \frac{\binom{2n}{n+i(k+1)}}{\binom{2n}{n}} = \sum_{i=-r}^{+r}(-1)^i \frac{\binom{2n}{n+i(k+1)}}{\binom{2n}{n}}.$$

さて $n \longrightarrow \infty$ のときの $P(D_{n,n} \leqslant k/n)$ を調べよう．便宜上，$\binom{2n}{n+i(k+1)}\Big/\binom{2n}{n}$ を $g_n(i,k)$ で表わし，$k = \lambda\sqrt{2n}$ とすると，任意の $\varepsilon > 0$ に対して，r' と $n_1(r',\varepsilon)$ を選んで，$n > n_1(r',\varepsilon)$ ならば

(14.3.80)
$$\left|2\sum_{i=r'+1}^{\infty}(-1)^i e^{-2i^2\lambda^2}\right| < \frac{\varepsilon}{3}$$

(14.3.81)
$$\left|P\left(D_{n,n} < \sqrt{\frac{2}{n}}\lambda\right) - \sum_{i=-r'}^{+r'}(-1)^i g_n(i,\lambda\sqrt{2n})\right| < \frac{\varepsilon}{3}$$

が成立するようにできる．(14.3.70) のように，$g_n(i,\lambda\sqrt{2n})$, $i = 0, \pm1, \cdots, \pm r'$ について大きな階乗に対するスターリングの公式を適用すれば，$n_2(r',\varepsilon)$ が存在して，$n > n_2(r',\varepsilon)$ ならば

(14.3.82)
$$\left|\sum_{i=-r'}^{+r'}(-1)^i g_n(i,\lambda\sqrt{2n}) - \sum_{i=-r'}^{+r'}(-1)^i e^{-2i^2\lambda^2}\right| < \frac{\varepsilon}{3}$$

が成り立つ．

(14.3.80), (14.3.81), (14.3.82) より $n > \max(n_1, n_2)$ のとき

(14.3.83)
$$\left|P(\sqrt{n/2}\,D_{n,n} < \lambda) - \sum_{i=-\infty}^{\infty}(-1)^i e^{-2i^2\lambda^2}\right| < \varepsilon$$

である．

要約すれば次の結果になる．

14.3.8　$F_{1n}(x)$, $F_{2n}(x)$ を同一で連続な c.d.f. からの大きさ n の独立な標本における経験 c.d.f. とすれば

(14.3.84)
$$P\left(D_{n,n} \leqslant \frac{k}{n}\right) = \sum_{i=-r}^{+r}(-1)^i \binom{2n}{n+i(k+1)}\Big/\binom{2n}{n}.$$

ただし r は $n/(k+1)$ 以下の最大整数．さらに

(14.3.85)
$$\lim_{n\to\infty} P(\sqrt{n/2}\,D_{n,n} < \lambda) = \sum_{i=-\infty}^{\infty}(-1)^i e^{-2i^2\lambda^2}.$$

ここに $D_{n,n}$ は (14.3.55) で定義されている．

もちろん読者は (14.3.85) が本質的には標本の大きさが n に等しいときの (14.3.57) の特別の場合にすぎないことに気づくだろう．D_n を大きさ n の1標本に対して (11.7.1) で定義すれば，$P(\sqrt{n}\,D_n < \lambda)$ は (14.3.85) と同じ極限を持つ．

(f)　マン=ホワイトニィ（ウィルコクソン）検定

ここでは重要な2標本問題を議論しよう．これは Wilcoxon (1945) が最初に，標本の大きさが等しい場合を考え，Mann と Whitney (1947) が大きさが等しくない場合へ拡張して，大標本に対する漸近的な結果も得ている．

$O_{n_1}: (x_1, \cdots, x_{n_1})$, $O_{n_2}: (x_1', \cdots, x_{n_2}')$ はそれぞれ連続な c.d.f. $F_1(x), F_2(x)$ からの標本とする．$z_{\xi\eta}$, $\xi = 1, \cdots, n_1$, $\eta = 1, \cdots, n_2$ は

(14.3.86)
$$z_{\xi\eta} = \begin{cases} 1, & x_\xi < x_\eta' \\ 0, & x_\xi \geqslant x_\eta' \end{cases}$$

で定義される確率変数の組である．いま

(14.3.87)
$$U = \sum_{\eta=1}^{n_2}\sum_{\xi=1}^{n_1} z_{\xi\eta}$$

として，$(x_1, \cdots, x_{n_1})$，$(x'_1, \cdots, x'_{n_2})$ を大きい順序に1列に並べれば，U は x が x' より先に位置する全回数を示している．この列の x と x' のすべてに，1から n_1+n_2 までの順位を割り当てたとき，x の順位の和 T をウィルコクソン統計量という．

(14.3.88)
$$U + T = n_1 n_2 + \frac{n_1(n_1+1)}{2}$$

が確かめられる．ウィルコクソンは実際は標本の大きさが等しいときだけ T を考えている．

マン=ホワイトニィ検定は U 統計量に基づく片側検定であり，大きさ α の棄却域 W_α が $U = 0, 1, \cdots, U_\alpha$ となる2標本の標本空間における事象からなるように構成されている．ただし U_α は

(14.3.89)
$$P(U \leqq U_\alpha \mid (F_1, F_2) \in \mathscr{J}_0) \leqq \alpha$$

を満たす最大整数．

$(F_1, F_2) \in \mathscr{J}_0$ を仮定して，$p_{n_1 n_2}(U)$ を U の p.f. とすれば，$p_{n_1 n_2}(U)$ は差分方程式

(14.3.90)
$$p_{n_1 n_2}(U) = \frac{n_1}{n_1+n_2} p_{n_1-1\, n_2}(U) + \frac{n_2}{n_1+n_2} p_{n_1 n_2-1}(U - n_1)$$

を満たす．ただし $U=0$ のとき $P_{oi}(U) = p_{io}(U) = 1$，$U \neq 0$ のとき $P_{oi}(U) = p_{io}(U) = 0$．この差分方程式を用いて，Mann と Whitney (1947) は $1 \leqq n_1 \leqq n_2 \leqq 8$ について $p_{n_1 n_2}(U)$ の値を作表している．

さて U の平均と分散を考えよう．マン=ホワイトニィ検定の一致性を調べるために，任意の2つの連続な c.d.f. $F_1(x), F_2(x)$，すなわち $(F_1, F_2) \in \mathscr{J}$ に関する U の平均と分散を求めると便利である．次が成り立つ．

(14.3.91)
$$\mathscr{E}(U) = \mathscr{E}\left(\sum_{\xi,\eta} z_{\xi\eta}\right) = \sum_{\xi,\eta} \mathscr{E}(z_{\xi\eta}) = n_1 n_2 a.$$

ただし

(14.3.92)
$$a = \int_{-\infty}^{\infty} F_1(x)\, dF_2(x).$$

また

(14.3.93)
$$\mathscr{E}(U^2) = \mathscr{E}\left(\sum_{\xi,\eta} z_{\xi\eta}\right)^2$$
$$= \sum_{\xi,\eta} \mathscr{E}(z_{\xi\eta}^2) + \sum_{\substack{\xi,\eta,\zeta \\ \xi\neq\eta}} \mathscr{E}(z_{\xi\zeta} z_{\eta\zeta}) + \sum_{\substack{\xi,\eta,\zeta \\ \eta\neq\zeta}} \mathscr{E}(z_{\xi\eta} z_{\xi\zeta}) + \sum_{\substack{\xi\neq\zeta \\ \eta\neq\omega}} \mathscr{E}(z_{\xi\eta} z_{\zeta\omega})$$

$$= n_1 n_2 a + n_1 n_2 (n_1 - 1) b + n_1 n_2 (n_2 - 1) c + n_1 n_2 (n_1 - 1)(n_2 - 1) a^2.$$

ただし

(14.3.94)
$$b = \int_{-\infty}^{\infty} (F_1(x))^2\, dF_2(x), \qquad c = \int_{-\infty}^{\infty} (1 - F_2(x))^2\, dF_1(x).$$

よって

(14.3.95)
$$\sigma^2(U) = n_1 n_2 [a + (n_1 - 1) b + (n_2 - 1) c - (n_1 + n_2 - 1) a^2].$$

$(F_1, F_2) \in \mathscr{J}_0$ ならば，$a = \frac{1}{2}$，$b = c = \frac{1}{3}$ で

(14.3.96)
$$\mathscr{E}(U) = \frac{n_1 n_2}{2}, \qquad \sigma^2(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}.$$

マンとホワイトニィは $(F_1, F_2) \in \mathscr{J}_0$ ならば，n_1，n_2 が大きいとき U が漸近的に $N(n_1 n_2/2, n_1 n_2 (n_1 + n_2 + 1)/12)$ に従って分布することを示している．

比 $U/n_1 n_2$ については

(14.3.97)
$$\mathscr{E}\left(\frac{U}{n_1 n_2}\middle| \mathscr{J}_0\right) = \frac{1}{2}, \qquad \sigma^2\left(\frac{U}{n_1 n_2}\middle| \mathscr{J}_0\right) = \frac{n_1 + n_2 + 1}{12 n_1 n_2}.$$

よって，チェビシェフの不等式 (3.3.5) より，$n_1, n_2 \longrightarrow \infty$ とき，$U/n_1 n_2$ は定数 $\frac{1}{2}$ に確率収束することがわかる．また $U_\alpha/n_1 n_2$ は $\frac{1}{2}$ に確率収束する．

さて U 検定の一致性を調べよう．$\mathscr{J}^-$ を

(14.3.98)
$$a < \frac{1}{2}$$

となる対 $(F_1(x), F_2(x))$ のクラスとすれば，任意の $(F_1(x), F_2(x)) \in \mathscr{J}^-$ に対して，$b < \frac{1}{3}$，$c < \frac{1}{3}$ が成り立つ．$\mathscr{J}^-$ の c.d.f. の任意の対に関する $U/n_1 n_2$ の平均と分散は

(14.3.99)
$$\mathscr{E}\left(\frac{U}{n_1 n_2}\right) = a < \frac{1}{2}$$
$$\sigma^2\left(\frac{U}{n_1 n_2}\right) = \frac{a + (n_1 - 1) b + (n_2 - 1) c - (n_1 + n_2 - 1) a^2}{n_1 n_2}$$

で与えられる．よって $n_1, n_2 \longrightarrow \infty$ のとき，$U/n_1 n_2$ が定数 $a < \frac{1}{2}$ に確率収束することは明らかである．したがって

(14.3.100)
$$\lim_{n_1, n_2 \to \infty} P\left(\frac{U}{n_1 n_2} < \frac{U_\alpha}{n_1 n_2}\middle| (F_1, F_2) \in \mathscr{J}^-\right) = 1.$$

すなわち

14.3.9 マン=ホワイトニィ検定は仮説 $\mathscr{H}(\mathscr{J}_0;\mathscr{J}_0\cup\mathscr{J}^-)$ すなわち任意の $(F_1,F_2)\in\mathscr{J}^-$ に対する $(F_1,F_2)\in\mathscr{J}_0$ の一致検定である.

$\mathscr{J}^+$ を $a>\frac{1}{2}$ となる対 (F_1,F_2) のクラスとすれば, F_1 と F_2 を入れ換えれば, この場合も上述の場合に帰着する.

いま, $U_{\frac{1}{2}\alpha}$ を $P(U\leqslant U_{\frac{1}{2}\alpha}|(F_1,F_2)\in\mathscr{J}_0)\leqslant\frac{1}{2}\alpha$ を満たす最大整数, $U'_{\frac{1}{2}\alpha}$ を $P(U\geqslant U'_{\frac{1}{2}\alpha}|(F_1,F_2)\in\mathscr{J}_0)\leqslant\frac{1}{2}\alpha$ を満たす最小整数とする. マン=ホワイトニィ検定を整数 $0,1,\cdots,U_{\frac{1}{2}\alpha}$ および $U'_{\frac{1}{2}\alpha},\cdots,n_1n_2$ からなる棄却域 W'_α を持つ両側検定として定義すれば, W'_α は任意の $(F_1,F_2)\in\mathscr{J}-\mathscr{J}_M$ に対する $(F_1,F_2)\in\mathscr{J}_M$ の一致検定になる. ただし $\mathscr{J}_M$ は $a=\frac{1}{2}$ なる連続な c.d.f. のすべての対からなるクラスである. $\mathscr{J}_M$ は同一で連続な c.d.f. の対のクラスではない. しかし $\mathscr{J}_0\subset\mathscr{J}_M\subset\mathscr{J}$ である.

14.4 確率化の方法

本節では, ある種のノンパラメトリック統計的仮説の検定を構成するために, 確率化の方法として知られている手法を考察しよう. 2種類の確率化検定 : **成分確率化検定**と**順位確率化検定**がある. これらを簡単に紹介しよう. 詳細に興味を持つ読者は Fraser (1957), Hoeffding (1948 *a*), Lehmann (1959), Lehmann と Stein (1949) を参照されたい.

(a) 成分確率化

成分確率化の方法は Fisher (1926 *b*) が最初に提唱した. この方法の原理は簡単であるが, 極小標本の場合を除けば個々の問題に適用するには技術的な困難が伴う. 標本空間として標本 $(x_1,\cdots,x_n)$ 成分を置換して得られる $n!$ 個の点を考えるのが, この方法の1標本に対する場合の最も簡単な型であろう.

一般に $(x_1,\cdots,x_n)$ が p.d.f. $f(x_1,\cdots,x_n)$ を持つとしよう. 確率0の集合における点を除けば, $(x_1,\cdots,x_n)$ の標本空間 R_n 内に, 与えられた標本成分の置換を座標に持つ $n!$ 個の異なる点が存在する. この $n!$ 個の点の集合を S で表わそう. すると $f_0(x_1,\cdots,x_n)$ が $x_1,\cdots,x_n$ に関して対称な p.d.f. ならば, 点 $(x_{\xi_1},\cdots,x_{\xi_n})$ に割り当てられる条件つき確率は

$$p_0(x_{\xi_1},\cdots,x_{\xi_n}|S)=\frac{f_0(x_{\xi_1},\cdots,x_{\xi_n})}{\sum^* f_0(x_{\xi_1},\cdots,x_{\xi_n})}=\frac{1}{n!} \tag{14.4.1}$$

で与えられる. ただし $\xi_1,\cdots,\xi_n$ は $1,\cdots,n$ の置換, $\sum^*$ は $1,\cdots,n$ の置換 $\xi_1,\cdots,\xi_n$ のすべてにわたる和である.

$f_1(x_1,\cdots,x_n)$ が対立する p.d.f ならば, $x_{\xi_1},\cdots,x_{\xi_n}$ に割り当てられる確率は

$$p_1(x_{\xi_1},\cdots,x_{\xi_n}|S)=\frac{f_1(x_{\xi_1},\cdots,x_{\xi_n})}{\sum^* f_1(x_{\xi_1},\cdots,x_{\xi_n})}. \tag{14.4.2}$$

$f_1(x_1,\cdots,x_n)$ に対する $f_0(x_1,\cdots,x_n)$ の最強力検定となる大きさ α の検定 W_α が存在して, 簡単のため α が $1/n!$ の倍数に選ばれるならば, W_α は $P(W_\alpha|f_1)$ が最大になるような S の $n!\alpha$ 個の点の集合でなければならない. S の $n!$ 個の点における p_1 の値がすべて異なるならば, このような検定は存在して一意である. 一般に, S に唯一の点が存在して, この点の p_1 の値が S の $n!\alpha-1$ 個の点における p_1 の値より大きくなく, S の他のすべての点における p_1 の値より大きいならば, このような検定が一意に存在する. この検定 W_α が効率的であるためには, どんな標本点 $(x_1,\cdots,x_n)$ が起っても, (ただし標本成分はすべて異なる) f_1 と f_0 を判別する検出力を持たねばならない.

$f_1(x_1,\cdots,x_n)$ も $x_1,\cdots,x_n$ に関して対称ならば, W_α は $f_1(x_1,\cdots,x_n)$ に対する $f_0(x_1,\cdots,x_n)$ の検定に対して検定力を持たない.

ここでは簡単のために唯一の標本 $(x_1,\cdots,x_n)$ の場合について成分確率化の概念を議論してきた. この考え方の最も興味ある重要な拡張は, 2標本あるいは多標本, 実験計画法, 多次元標本などの問題についてである. しかしこれらの問題に成分確率化の方法を厳密に適用しようとすれば, 極小標本のとき以外はかなり技術的な困難が生じる. 本質的な困難は帰無仮説に対する f_1 の下で, p_1 の標本成分の置換上の値を評価して, 検定 W_α が条件を満たすように S 内の $\alpha n!$ 個の点を選ぶことである. ノンパラメトリック検定を構成する際に成分確率化の方法を用いてきた人々は発展させるために成分確率化の原則を厳密に適用することをゆるめて, パラメトリック検定論から検定関数 $g(x_1,\cdots,x_n)$ を借りて棄却域 W_α を定めた. この検定関数の標本成分の置換全体の空間 S にわたるモーメントおよび分布関数の大標本の場合の性質などが, 帰無仮説の下あるいは時として対立仮説の下において調べられている. 置換全体の空間 S における棄却域 W_α は g の値がある区間あるいは集合にはいるような置換点の集合になる.

これまで開発された有用な成分確率化検定をすべて考察することは不可能である. そこ

で次に示す2つの例を考えれば十分だろう．

(b) 成分確率化検定の例

最初の成分確率化検定は**フィッシャー＝ピットマン検定**として知られている2標本検定である．これは，2標本が平均の異なる分布から抽出されたという対立仮説に対して，2標本が同一の平均を持つ分布から抽出されたという仮説の検定を意図している．Fisher (1926 *b*) は最初に個々の問題についてこの検定を取り扱い，一般化は Pitman (1937 *b*) によりなされた．

$O_{n_1}:(x_1,\cdots,x_{n_1})$ と $O_{n_2}:(x'_1,\cdots,x'_{n_2})$ をそれぞれ連続な c.d.f. からの独立な2標本とする．この2つの標本を1つに合わせた n_1+n_2 個の成分の任意の置換を $y_1,\cdots,y_{n_1+n_2}$ で表わそう．$\bar{y}$, s^2 を $(y_1,\cdots,y_{n_1})$ の任意の標本平均，標本分散とし，$\bar{y}'$, s'^2 を $(y_{n_1+1},\cdots,y_{n_1+n_2})$ の標本平均，標本分散とする．検定関数

$$(14.4.3)\qquad g(y_1,\cdots,y_{n_1+n_2})=\frac{n_1n_2(\bar{y}-\bar{y}')^2}{(n_1+n_2)[(n_1-1)s^2+(n_2-1)s'^2]+n_1n_2(\bar{y}-\bar{y}')^2}$$

が考えられている．この $g(y_1,\cdots,y_{n_1+n_2})$ は $t^2/(n_1+n_2-2+t^2)$ なる型をしている．ただし t は，《2標本が等分散を持つ2つの正規分布から抽出された条件の下で，同一の正規分布である》という仮説に対するスチューデントの t 比である．

ピットマンは $y_1,\cdots,y_{n_1+n_2}$ が与えられたときの $g(y_1,\cdots,y_{n_1+n_2})$ の条件つき分布の，n_1+n_2 個の y のすべての置換の集合上での，最初の4つのモーメントを求めている．g の分母は y のすべての置換上で定数だから g の分子だけを考えればよい．この分子は n_1+n_2 個の標本成分を n_1 個の成分と n_2 個の成分からなる2つの集合に分ける異なる方法の数に対応して，高々 $\binom{n_1+n_2}{n_1}$ 個の異なる値を持つ．この置換上の g の分布は漸近的にベータ分布 $Be\left(\frac{1}{2},\frac{1}{2}(n_1+n_2-2)\right)$ に従って分布することがピットマンにより示されている．最後に棄却域 W_α は $P(g>g_\alpha)=\alpha$ なる g の値を与えるすべての置換からなることに注意しよう．ただし g_α はベータ分布 $Be\left(\frac{1}{2},\frac{1}{2}(n_1+n_2-2)\right)$ の右側 100α% 点として近似される．O_{n_1} と O_{n_2} の平均間の差は $g(x_1,\cdots,x_{n_1},x'_1,\cdots,x'_{n_2})>g_\alpha$ ならば 100α% 有意水準で有意であると判定される．Wald と Wolfowitz (1944) は大きな n_1 と n_2 に対する g の漸近分布を考察して，n_1+n_2 の値が大きいとき，ベータ分布 $Be\left(\frac{1}{2},\frac{1}{2}(n_1+n_2-2)\right)$ で与えられた g_α は漸近的に正しいことを示すことと同値になる証明を与えている．

成分確率化検定の2番目の例として標準型モデル I 分散分析検定を考えよう．すなわち，r 個の行と s 個の列の実験計画配置において各行の観測が1つの連続な c.d.f. を持つ母集団からの大きさ s の標本を構成し，列に対応する c.d.f. は平均以外は同一であるという仮定の下で，列効果がすべて0という仮説検定を考える [10.6節を見よ]．$(x_{\xi\eta},\ \xi=1,\cdots,r,\ \eta=1,\cdots,s)$ が標本のとき Welch (1937) および Pitman (1938) は検定関数

$$(14.4.4)\qquad g(\{x_{\xi\eta}\})=\frac{\sum_{\xi,\eta}(\bar{x}_{\cdot\eta}-\bar{x})^2}{\sum_{\xi,\eta}(x_{\xi\eta}-\bar{x}_{\cdot\eta}-\bar{x}_{\xi\cdot}+\bar{x})^2+\sum_{\xi,\eta}(\bar{x}_{\cdot\eta}-\bar{x})^2}$$

を用いた．ただし $\sum_{\xi,\eta}$ は $\xi=1,\cdots,r,\ \eta=1,\cdots,s$ に関する和，$\bar{x}_{\xi\cdot}$, $\bar{x}_{\cdot\eta}$ は $\|x_{\xi\eta}\|$ におけるそれぞれ ξ 番目の行平均および η 番目の列平均，$\bar{x}$ は rs 個の x すべての標本平均である．ピットマンは成分確率化の方法を用いて，x の行内の置換すべての下で g の最初の4つのモーメントを計算し，この分布がベータ分布 $Be\left(\frac{1}{2}(s-1),\frac{1}{2}(r-1)(s-1)\right)$ で十分に近似され得ることを示した．このベータ分布はすでに 10.6 (a) 節で述べたように，列効果0を仮定したときの正規理論における g の分布に他ならない．この検定の置換全体の空間における点の棄却域 W_α は $P(g>g_\alpha)=\alpha$ なるすべての置換からなる．ただし g_α は上述のベータ分布の右側 100α% 点である．

Arnold (1958) はこの確率比検定を $x_{\xi\eta}$ がベクトルで $s=2$ の場合に拡張している．この場合，検定は第18章で定義して論じられる2標本に対するホテリングの T^2 と同値である．Welch (1937) も成分確率化の方法を用いて，ラテン方格に対するモデル I 分散分析を調べている．

Hoeffding (1952) は上述の2つの例を含む大標本に対する成分確率化検定の検出力について研究し，その検出力は大標本に対しては，対応するパラメトリック検定の場合と漸近的に等しいことを示している．

(c) 順位確率化検定

成分確率化の方法は，厳密には分布によらない手法とはいえない．なぜなら，成分確率化の下では検定関数およびその分布論は観測あるいは実現される標本の成分値に依存するからである．しかし順位確率化検定を用いれば，分布によらないという性質は保たれる．順位確率化検定では標本成分自身よりもむしろ標本成分の順位が用いられる．$x_1,\cdots,x_n$ を連続な c.d.f. $F(x_1,\cdots,x_n)$ からの n 次元確率変数の成分とする．$x_1,\cdots,x_n$ の順位

$a_1, \cdots, a_n$ は x を増加する順に並べたときの x の右下の添字で表わされる．すなわち a_ξ は $(x_1, \cdots, x_n)$ の ξ 番目に小さい x の添字である．よって $a_1, \cdots, a_n$ は $1, \cdots, n$ の1つの置換である．$1, \cdots, n$ の $n!$ 個の置換すべての空間を S' としよう．$(x_1, \cdots, x_n)$ の標本空間 R_n の $x_{\xi_1} < \cdots < x_{\xi_n}$ なる点はすべて空間 S' の点 $(\xi_1, \cdots, \xi_n)$ に写像される．$(x_1, \cdots, x_n)$ が x に関して対称な c.d.f. $F_0(x_1, \cdots, x_n)$ を持つならば，S' の各点で

$$p_0(\xi_1, \cdots, \xi_n) = \frac{1}{n!} \tag{14.4.5}$$

である．(14.4.1) より，読者は $p_0(x_{\xi_1}, \cdots, x_{\xi_n} | S)$ が $(x_1, \cdots, x_n)$ を固定した条件つき確率であることに気づくであろう．しかし対応する順位の分布 $p_0(\xi_1, \cdots, \xi_n)$ は条件つき分布ではない．$F_1(x_1, \cdots, x_n)$ が任意の連続な n 次元 c.d.f. ならば，確率分布 $p_1(\xi_1, \cdots, \xi_n)$ は

$$p_1(\xi_1, \cdots, \xi_n) = \int_{E_n} dF_1(x_1, \cdots, x_n) \tag{14.4.6}$$

で与えられる．ただし E_n は $x_{\xi_1} < \cdots < x_{\xi_n}$ なる領域である．

c.d.f. F_1 に対して c.d.f. F_0 を検定するために，S' から棄却域 W'_α を構成する問題は，すでに述べたように，p.d.f. f_1 に対する p.d.f. f_0 の検定のため S から棄却域 W_α を構成するのとまったく同様である．しかし W'_α の場合は W_α のときのように検定は観測された点 $(x_1, \cdots, x_n)$ に依存しない．この意味で W'_α は W_α より真にノンパラメトリックである．c.d.f. F_0, F_1 がともに $x_1, \cdots, x_n$ に関して対称ならば，順位確率化検定は F_0 と F_1 を判別する検出力を持たない．

成分確率化の場合と同様に，W'_α を定めるのに順位確率化の原則を厳密に適用しようとすれば，技術的に相当困難になるだろう．順位確率化検定論をさらに展開するためには，類似したパラメトリック検定問題で提案されたように，S' のすべての点で定義された検定関数 $h(a_1, \cdots, a_n)$ を用いる必要があると指摘されている．

相当数の順位検定が展開されているが，次にそれを例で3つだけ考察しよう．

(d) 順位確率化検定の例

(14.3.88) のウィルコクソン検定 T は仮説 $\mathscr{H}(\mathscr{A}; \mathscr{B})$ の検定に対する順位検定である．ただし，$\mathscr{A}$ (帰無仮説) は連続で対称な $F(x_1) \cdots F(x_{n_1}) F(x'_1) \cdots F(x'_{n_2})$ なる型をした $(n_1 + n_2)$ 次元 c.d.f. のクラスである．一方，$\mathscr{B}$ は連続な $F_1(x_1) \cdots F_1(x_{n_1}) F_2(x'_1) \cdots F_2(x'_{n_2})$ なる型で，$a < \frac{1}{2}$ (a は (14.3.92) で定義されている) なる $(n_1 + n_2)$ 次元 c.d.f. クラスである．$x_1, \cdots, x_{n_1}, x'_1, \cdots, x'_{n_2}$ を $y_1, \cdots, y_{n_1+n_2}$ として，y の順位を $a_1, \cdots, a_{n_1+n_2}$ で表わせば，ウィルコクソン統計量 T は単に $h(a_1, \cdots, a_{n_1+n_2})$ となる．ただし

$$h(a_1, \cdots, a_{n_1+n_2}) = \sum_1^{n_1+n_2} a_i. \tag{14.4.7}$$

順位統計量に基づく確率化検定のもう1つの例として，**順位相関係数**がある．$(x_1, y_1; x_2, y_2; \cdots; x_n, y_n)$ を連続な c.d.f. $F(x, y)$ を持つ母集団からの大きさ n の標本とする．$a_1, \cdots, a_n$ を $x_1, \cdots, x_n$ の順位，$b_1, \cdots, b_n$ を $y_1, \cdots, y_n$ の順位とする．ここで用いられる順位統計量 $g(a_1, \cdots, a_n, b_1, \cdots, b_n)$ は順位の2つの集合間の相関係数 R である．R によって検定される仮説は $\mathscr{H}(\mathscr{A}; \mathscr{B})$ である．ただし $\mathscr{B}$ は $F(x_1, y_1) \cdots F(x_n, y_n)$ なる型の連続な $2n$ 次元 c.d.f. のクラス，$\mathscr{A}$ (帰無仮説) は $F(x_\xi, y_\xi) = F_1(x_\xi) F_2(y_\xi)$, $\xi = 1, \cdots, n$ で $F_1(x)$ と $F_2(x)$ が連続な c.d.f. であるような $\mathscr{B}$ の部分クラスである．Hotelling と Pabst (1936) は，帰無仮説が正しいとき，すなわち $(x_1, y_1; \cdots; x_n, y_n)$ の分布が $\mathscr{A}$ に属すとき，R の分布を調べ，大きな n に対して $\sqrt{n} R$ は漸近的に $N(0, 1)$ に従って分布することを示している．Olds (1938) は R と関係式

$$R = 1 - 6S/(n^3 - n) \tag{14.4.8}$$

で結ばれる S の分布表を作成している．

検定 R の棄却域 W_α が $R^2 > R_\alpha^2$ なる標本空間内の集合のとき，Hoeffding (1948 *b*) は，この検定が $\mathscr{B} - \mathscr{A}$ の任意の c.d.f. に対する $\mathscr{A}$ の任意の c.d.f. の一致検定でないことを示した．ただし R_α^2 は $P(R^2 \geqslant R_\alpha^2) \leqslant \alpha$ なる最小数である．

Hoeffding (1948 *b*) はすでに定義された $\mathscr{H}(\mathscr{A}; \mathscr{B})$ の検定に対する順位に基づく確率化検定方式を開発している．その統計量は

$$D = \frac{(n-5)!}{n!}[A - 2(n-2)B + (n-2)(n-3)C] \tag{14.4.9}$$

である．ただし

$$\begin{aligned} A &= \sum_{\xi=1}^n a_\xi b_\xi (a_\xi - 1)(b_\xi - 1) \\ B &= \sum_{\xi=1}^n (a_\xi - 1)(b_\xi - 1) c_\xi \\ C &= \sum_{\xi=1}^n c_\xi (c_\xi - 1) \end{aligned} \tag{14.4.10}$$

で，c_ξ は $x_\eta < x_\xi$ かつ $y_\eta < y_\xi$ なる標本の対 (x_ξ, y_ξ), (x_η, y_η) の数である．

D は

(14.4.11) $$\Delta = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[F(x,y) - F(x,\infty)F(\infty,y)]^2\,dF(x,y)$$

に対する不偏推定量で，帰無仮説の下でのこの分散は

(14.4.12) $$\sigma^2(D) = \frac{2(n^2+5n-32)}{8100n(n-1)(n-2)(n-3)}$$

で与えられる．ホエフディングは $2n$ 次元分布が $\mathscr{B}-\mathscr{A}$ に属すという仮説の下で，$\sigma^2(D)$ を求めている．

分散分析における種々の仮説検定に対する重要な順位確率化検定方式が Friedman (1937), Kendall と Smith (1939), Kruskal (1952), Wallis (1939) により開発されている．Wormleighton (1959), Andrews (1954) は大標本におけるこれらの検定の検出力について研究している．

問　題

14.1 $\mathscr{C}_0$ をそれぞれ $x_{01}, x_{02}, \cdots, x_{0k}$ で $p_1, p_2, \cdots, p_k$ 分位を持つ連続な c.d.f. すべてのクラスとする．ただし $0 = p_0 < p_1 < \cdots < p_k < p_{k+1} = 1$. $\mathscr{C}$ をすべての連続な c.d.f. のクラスとする．$I_1, \cdots, I_{k+1}$ はそれぞれ区間 $(-\infty, x_{01}], (x_{01}, x_{02}], \cdots, (x_{0k}, +\infty)$ である．大きさ n の標本に対して，$n_1, \cdots, n_{k+1}$, $(n_1 + \cdots + n_{k+1} = n)$ をそれぞれ $I_1, \cdots, I_{k+1}$ にはいる標本成分の数としよう．大きな n のとき

$$\sum_{i=1}^{k+1}\frac{[n_i - n(p_i - p_{i-1})]^2}{n(p_i - p_{i-1})} > \chi^2_{k,\alpha}$$

を満たす $(n_1, \cdots, n_{k+1})$ の標本空間内の点集合 W_α は仮説 $\mathscr{H}(\mathscr{C}_0 ; \mathscr{C})$ の大きさ α の一致検定であることを示せ．ただし $P(\chi^2 > \chi^2_{k,\alpha}) = \alpha$ で，χ^2 はカイ2乗分布 $C(k)$ を持つ確率変数である．

14.2 $\mathscr{C}_{0p}$ を $F_1(x_{0p}) = \cdots = F_n(x_{0p}) = p$ を満たす $F_1(x_1)F_2(x_2)\cdots F_n(x_n)$ なる型の n 次元 c.d.f. すべてのクラスとする．$\mathscr{C}_p$ は $F_\xi(x_{0p}) = p_\xi$, $\xi = 1, \cdots, n$ を満たす $F_1(x_1)F_2(x_2)\cdots F_n(x_n)$ なる型の n 次元 c.d.f. すべてのクラスとし，$n \longrightarrow \infty$ のとき $(p_1 + \cdots + p_n)/n \longrightarrow p' \neq p$ とする．r を $(-\infty, x_{0p}]$ にはいる n 次元確率変数 $(x_1, \cdots, x_n)$ の成分個数とし，W_α を (14.1.2) で定義する．このとき，$n \longrightarrow \infty$ ならば，W_α は $\mathscr{H}(\mathscr{C}_{0p} ; \mathscr{C}_p)$ の大きさ α の一致検定になることを示せ．

14.3 (14.2.12) を確かめよ．

14.4 (14.2.12) より，$m, n \longrightarrow \infty$, $n/m \longrightarrow \rho > 0$ のとき

$$\lim \mathscr{E}\left(\frac{s_i}{m}\right) = \frac{\rho^i e^{-\rho}}{i\,!}$$

を示せ．

14.5 14.2 (c) 節で定義された確率変数 s_1 は $n = \rho m + O(1)$, $\rho > 0$ のとき，大きな m, n に対して，漸近的分布として

$$N(m\rho e^{-\rho}, m\rho e^{-\rho}(1 - e^{-\rho} - \rho e^{-\rho}))$$

を持つことを示せ．

14.6 (14.2.15) における $u_0^\delta v^n$ の係数は (14.2.16) における $u_0^\delta v^n$ の係数に一致することを示せ．

14.7 (14.2.24) を証明せよ．

14.8 $(r_1^*, \cdots, r_t^*)$ が p.f. (14.3.7) を持つ確率変数 $(r_1, \cdots, r_{n_1})$ の任意 t 個ならば，$(r_1^*, \cdots, r_t^*)$ の p.f. は (14.3.8) で与えられることを示せ．

14.9 (14.3.13) を確かめよ．

14.10 (14.3.14) を確かめよ．

14.11 (14.3.14) より，$n_1, n_2 \longrightarrow \infty$, $n_1/n_2 \longrightarrow \rho > 0$ のとき

$$\lim \mathscr{E}\left(\frac{s_i}{n_1+1}\right) = \frac{\rho^i}{(1+\rho)^{i+1}}$$

を示せ．

14.12（続き） 前問と同じ極限条件のとき

$$\lim \mathscr{E}\left(\frac{s_i}{n_1+1}\middle|\mathscr{J}^*\right) = \int_0^1 \frac{[\rho g(u)]^i}{[1+\rho g(u)]^{i+1}}\,du$$

を示せ．

14.13 14.3.1 において m_1 を z_k を超える x の個数，m_2 を z_k を超える x' の個数とすれば，(m_1, m_2) の p.f. が

$$p(m_1, m_2) = \frac{\binom{n_1}{m_1}\binom{n_2}{m_2}}{\binom{n_1+n_2}{k}}$$

で与えられることを示せ．ただし m_1, m_2 は $m_1 + m_2 = n_1 + n_2 - k$ を満たす非負整数 [Mood (1950)].

14.14 $\mathscr{J}_1$ を $(F(x), F(x-\delta))$ なる型の連続な c.d.f. の対全体のクラス（ただし δ は定数）とし，$\mathscr{J}_0$ を同一で連続な c.d.f. の対全体のクラスとする．マン=ホワイトニィ検定は $\delta > 0$ なる $\mathscr{J}_1$ の任意の要素に対して $\mathscr{J}_0$ の任意の要素を検定する一致検定であることを示せ．

14.15（続き） $(x_{(1)}, \cdots, x_{(n_1)})$, $(x'_{(1)}, \cdots, x'_{(n_2)})$ がそれぞれ $F(x)$, $F(x-\delta)$ からの大きさ n_1, n_2 の独立な標本の順序統計量であるとき

$$P(x_{(r)} - x'_{(s)} < \delta) = \sum_{i=1}^{r-1}\left(\frac{s}{s+i}\right)\binom{n_1}{i}\binom{n_2}{s}\Big/\binom{n_1+n_2}{i+s}$$

を示せ．よって $r < s$ かつ $r' > s'$ ならば，$(x'_{(s')} - x_{(r')}, x'_{(s)} - x_{(r)})$ は δ に対する $100\gamma\%$ 信頼区間である．ただし

$$\Upsilon = 1 - \sum_{i=1}^{r-1} \left(\frac{s}{s+i}\right)\binom{n_1}{i}\binom{n_2}{s}\bigg/\binom{n_1+n_2}{i+s} - \sum_{i=r'}^{n_1} \left(\frac{s'}{s'+i}\right)\binom{n_1}{i}\binom{n_2}{s'}\bigg/\binom{n_1+n_2}{i+s'}.$$

[Mood (1950)].

14.16 (14.3.78) を確かめよ.

14.17 (14.3.88) を確かめよ.

14.18 差分方程式 (14.3.90) を導け.

14.19 (14.3.93) を確かめよ.

14.20 $\mathscr{E}(D) = \Delta$ を示せ. ただし, D, Δ はそれぞれ (14.4.9), (14.4.11) で定義されている.

14.21 Bertrand (1887) の投票問題 投票箱に, A 候補に記入した用紙が N_1 枚, B 候補に投票した用紙が N_2 枚はいっているとする. ただし $N_1 > N_2$ である. 開票は一度に1枚ずつ行なわれるとする. i 枚の投票用紙が取り出されたときの A 候補の得票数から B 候補のそれを引いた数を y_i とする. 全部の投票用紙を取り出す $\binom{N_1+N_2}{N_1}$ 通りの異なる可能な順序が等確率ならば, (14.3 (e) 節で用いた方法によって) 開票中, 終始, A 候補が B 候補よりリードしている (すなわち, $y_1, \cdots, y_{N_1+N_2}$ がすべて >0 となる) 確率は $(N_1 - N_2)/(N_1 + N_2)$ であることを示せ.

14.22 (続き) $N_1 = N_2 = N$ を仮定する. このとき, $y_{2N} = 0$ である. xy 平面上の $2N+1$ 個の連続した点 $(0,0), (1, y_1), \cdots, (2N, 0)$ で生成される $(0,0)$ を始点, $(2N, 0)$ を終点とする $\binom{2N}{N}$ 通りの折れ線パスを考える. この場合, 任意のパスが x 軸より上 (あるいは下) に位置する部分の数は, つねに偶数である. $2n$ 個の部分が x 軸より上 ($2N - 2n$ 個が下) に位置するパスの数は

$$\binom{2N}{N}\bigg/(N+1)$$

であることを示せ. したがってすべてのパスが等確率ならば, パスが x 軸より n/N の割合で上に位置する確率は

$$P_N(n) = \frac{1}{N+1}, \qquad n = 0, 1, \cdots, N$$

で与えられる.

14.23 硬貨投げの長時間経過 [Chung と Feller (1949)] 硬貨を $2N$ 回投げて, y_i を最初の i 回の投げにおいて, 表の出た回数から裏の出た回数を引いた数とする. 点 $(0, 0), (1, y_1), \cdots, (2N, y_{2N})$ を結ぶ 2^{2N} 通りの折れ線パスを考える. $2n$ 個の部分が x 軸より下 (もちろん, $2N - 2n$ 個の部分が上) に位置するパスの数は

$$\binom{2n}{n}\binom{2N-2n}{N-n}$$

であることを示せ. したがってパスが x 軸より上に n/N の割合で位置する確率は, すべてのパスが等確率 (すなわち, $2N$ 回独立に正しい硬貨を投げるとする) ならば

$$\binom{2n}{n}\binom{2N-2n}{N-n}\bigg/2^{2N}, \qquad n = 0, 1, \cdots, N$$

であることを示せ.

14.24 (続き) **(最初の) 逆正弦法則** [Chung と Feller (1949)] 確率変数 n/N の, $N \longrightarrow \infty$ のときの極限 c.d.f. $G(u)$ は大きな階乗に対するスターリングの公式を用いれば

$$G(u) = \frac{2}{\pi} \sin^{-1} \sqrt{u}$$

で与えられることを示せ. この極限 c.d.f. の p.d.f. は $g(u) = 1/[\pi\sqrt{u(1-u)}]$ で与えられる. これは区間 $(0, 1)$ の端点で無限の密度を持つ U 型の密度関数である. よって, 比較的大きな (あるいは, 比較的小さな) 割合でパスが x 軸より上に位置する頻度は, 1/2 近くの割合で x 軸より上に位置するよりもはるかに大きい. これは直感的にも明らかであろう.

第15章　逐次統計解析

15.1　概　　説

第8章から第14章までに論じられた標本論，推定論および仮説検定においては標本の大きさは固定されていた．これらの理論には特定の統計量の標本分布，すなわち与えられた標本の大きさに対する標本要素の関数の標本分布を求めることが含まれていた．この分布のうちいくつかについては標本の大きさが限りなく大きくなるときの極限形が求められた．しかし第6章および第7章では，いくつかの待ち時間分布を議論した．この分布ではある目標に達するまでに要する試行の回数は確率変数である．このような待ち時間分布に見られる過程は単純な逐次的な統計的処理の例であろう．本章では主に統計的仮説検定を中心に論じよう．しかし逐次推定に関する結果もいくつか述べる．

ここでは単純仮説検定に貫かれている基本的な結果を考える．複合仮説検定に関する結果は Wald の著作 (1947 *a*) に見い出される．その後は Barnard (1952), Cox (1952) による．David と Kruskal (1956), Rushton (1950, 1952) には逐次 t 検定に関する結果が得られている．

2回抜取検査方式を開発した Dodge と Romig (1929) が最初に逐次検定を考えた．この基本的な考え方は未知の $N\theta$ 個の不良品を含んでいる N 個の製品のロットを合格とするか不合格とするかどうかを判定する枠組の設定にある．ただし θ は $[0,1]$ 上の $\dfrac{1}{N}$ の倍数である．ドッジとロミッグの枠組の下では，ロットから大きさ n_1 の標本を抽出して標本の中の不良品の個数 m_1 によって次のような3つの戦略的な決定を行なう．

(i)　$m_1 \leqslant c_1$ ならば，このロットを合格とする．

(ii)　$m_1 > c_2$ ならば，ロットを不合格とする．

(iii)　$c_1 < m_1 \leqslant c_2$ ならば，大きさ n_2 の2番目の標本を抽出する．

決定 (iii) を行なう場合，すなわち次に大きさ n_2 の標本を抽出するならば，この標本

に含まれる不良品の個数 m_2 に依存して，次の2つの決定のうち1つが下される．

(i)　$m_1 + m_2 \leqslant c_2$ ならば，ロットを合格とする．

(ii)　$m_1 + m_2 > c_2$ ならば，ロットを不合格とする．

2回抜取検査は4つの数 n_1, n_2, c_1, c_2 を与えることによって完全に定められる．全標本の大きさ n は明らかに n_1, $n_1 + n_2$ の2つの値を取り得る確率変数である．1回抜取検査方式および2回抜取検査方式に関しては Dodge と Romig (1959) で詳しく取り上げられている．

2回抜取検査方式がその決定機構とともに本質的に統計的検定を構成することをみるためには次の事柄に注意しなければならない．すなわちロットが不合格となるのは2次元確率変数 (m_1, m_2) が $E_1 \cup E_2$ で定義される棄却域 W にはいるときであり，かつこのときに限る．ただし，E_1 は (m_1, m_2) の標本空間上の $m_1 > c_2$ なる点集合で，E_2 は $c_1 < m_1 \leqslant c_2$ かつ $m_1 + m_2 > c_2$ なる点集合である．不良品の割合が θ であるときロットが不合格になる確率 $P(W|\theta)$ は超幾何分布で容易に表わされる．もちろん $P(W|\theta)$ は θ の関数とみれば，"検定" W の検出力関数である．いま

$$L(\theta) = P(\overline{W}|\theta)$$

とすれば $L(\theta)$ は不良率 θ のロットを合格とする確率である．13.1節の歴史的な注意で述べた用語を用いれば，次のようにいい換えられる．$\theta \leqslant \theta_0$ のとき，ロットは良好な領域にあり，$\theta_0 < \theta < \theta_1$ のとき無関心領域にあり，$\theta \geqslant \theta_1$ のとき不合格の領域にある．$\theta = \theta_0$ における生産者危険（第1種の誤りの危険）α および $\theta = \theta_1$ における消費者危険（第2種の誤りの危険）β を与えれば

$$L(\theta_0) = 1 - \alpha, \qquad L(\theta_1) = \beta$$

となる．ただし実際の適用例では，α, β は通常区間 $[0.01, 0.10]$ 内の値が用いられる．検査特性曲線，すなわち θ の関数として見た $L(\theta)$ のグラフは点 $(0,1)$, $(\theta_0, 1-\alpha)$, (θ_1, β), $(1, 0)$ を通る．

真に逐次的な過程の考え方は Bartky (1943) による．彼は2回抜取検査の概念を拡張して，未知の不良品率 θ を持つ無限に大きなロットについて多回抜取検査を考えた．他に多段過程を数学的に研究した初期の頃の例としては，ベンガルの黄麻面積に関する標本論の Mahalanobis (1940) 級数，回帰関数の最大値を求めるための実験列に関する Hotelling (1941) の考え方，および Dixon と Mood (1948) による確率近似法が含まれる．しかし今日知られているような逐次解析の一般論へと大きく進歩したのは Wald (1945)

に負うところが大きい．この章では逐次解析におけるワルドの基本的な結果のいくつかを簡単に示そう．古典的な結果に深く興味を持たれる読者は Wald (1947 a) の著作を参照されたい．逐次的な過程に関する最近の数学的な研究については Dvoretzky, Kiefe と Wolfowitz (1953 a, 1953 b) の論文，Kiefer (1948, 1953, 1957), Kiefer と Wolfowitz (1952), Robbins (1952), Robbins と Monro (1951) 等が参考になるだろう．

15.2 逐次検定の基本構造

(a) 逐次検定における事象の記述

すでに述べたドッジとロミッグの2回抜取検査の枠組は一般の逐次検定の本質的な構造を暗示している．実数のパラメータ θ に依存する確率過程 $(x_1, x_2, \cdots)$ があるとしよう．これは成分 $x_1, x_2, \cdots$ の任意有限個がパラメータ空間 Ω の θ に依存する c.d.f. を持っていることを意味している．このように $n=1,2,\cdots$ に対して $(x_1,\cdots,x_n)$ は c.d.f. $F(x_1,\cdots,x_n;\theta)$ を持っている．まず単純仮説 $\mathscr{H}(\theta_0;\Omega):\theta=\theta_0$（今後簡単に $\mathscr{H}_0$ と書く）を考えよう．$R_1^{(1)},\cdots,R_1^{(n)}$ はそれぞれ $x_1,\cdots,x_n$ の標本空間で $R_n=R_1^{(1)}\times\cdots\times R_1^{(n)}$ $(n=1,2,\cdots)$ を $(x_1,\cdots,x_n)$ の標本空間とする．R_∞ として $R_\infty=R_1^{(1)}\times R_1^{(2)}\times\cdots$ を用いると便利である．

各正整数 n について，G_n°, G_n', G_n を次のような R_n における排反な事象とする．

(15.2.1) $$G_n^\circ\cup G_n'\cup G_n=G_{n-1}\times R_1^{(n)}.$$

まず x_1 を抜き取って逐次実験を始める．実験の停止，および続行に対しては次の規則を用いる．

(15.2.2)
$x_1\in G_1^\circ$ ならば，次を抜き取らずに $\mathscr{H}_0$ を採択する．
$x_1\in G_1'$ ならば，次を抜き取らずに $\mathscr{H}_0$ を棄却する．
$x_1\in G_1$ ならば，x_2 を抜き取る．

そして

(15.2.3)
$(x_1,x_2)\in G_2^\circ$ ならば，次を抜き取らずに $\mathscr{H}_0$ を採択する．
$(x_1,x_2)\in G_2'$ ならば，次を抜き取らずに $\mathscr{H}_0$ を棄却する．
$(x_1,x_2)\in G_2$ ならば，x_3 を抜き取る．

さらに一般的には，$n=1,2,\cdots$ に対して

(15.2.4)
$(x_1,\cdots,x_n)\in G_n^\circ$ ならば，次を抜き取らずに $\mathscr{H}_0$ を採択する．
$(x_1,\cdots,x_n)\in G_n'$ ならば，次を抜き取らずに $\mathscr{H}_0$ を棄却する．
$(x_1,\cdots,x_n)\in G_n$ ならば，x_{n+1} を抜き取る．

便宜上，$G_1, G_2,\cdots$ を**実験続行事象**と呼ぼう．

$G_1^\circ,\cdots,G_{n-1}^\circ$, $G_1',\cdots,G_{n-1}'$ は R_n における筒事象であり，G_n°, G_n', G_n は R_n の事象で，その $(2n+1)$ 個の事象は排反で

(15.2.5) $$G_1^\circ\cup\cdots\cup G_n^\circ\cup G_1'\cup\cdots\cup G_n'\cup G_n=R_n$$

である．しかも任意の正整数 n に対してこの同じ $(2n+1)$ 個の事象は R_∞ における排反な筒事象で，n を動かしたときのこの事象の和は R_∞ 自身になる．

事象 $G_1^\circ,\cdots,G_n^\circ$, $G_1',\cdots,G_n'$, G_n, $n=1,2,\cdots$ は (15.2.2), (15.2.3) および (15.2.4) で定義された決定過程とともに $\mathscr{H}_0$ に対する逐次検定 S を定義している．$\mathscr{H}_0$ の棄却域 W_α は $P(W_\alpha|\theta_0)=\alpha$ となるように G の項を選んだときの $G_1'\cup G_2'\cup\cdots$ である．1段ずつ逐次検定を行なう過程を S に対する**逐次過程**という．

標本の大きさを n に固定した標本空間において棄却域 W を持つ $\mathscr{H}_0$ に対する任意の検定は $G_1^\circ,\cdots,G_{n-1}^\circ$, $G_1',\cdots,G_{n-1}'$, G_n を空集合，$G_n'=W$, $G_n^\circ=\overline{W}$ とした**退化逐次検定**と呼ばれるものになる．

(b) 逐次検定における確率

x_n の抽出に基づいて $\mathscr{H}_0$ を採択する確率，$\mathscr{H}_0$ を棄却する確率，および x_{n+1} を抜き取る確率はそれぞれ

(15.2.6) $$P(G_n^\circ|\theta),\qquad P(G_n'|\theta),\qquad P(G_n|\theta)$$

である．すでに述べたように事象 $G_1^\circ,\cdots,G_n^\circ$, $G_1',\cdots,G_n'$, G_n は排反で，その確率はパラメータ θ の関数である．

(15.2.7)
$$L_n(\theta)=P(G_1^\circ|\theta)+\cdots+P(G_n^\circ|\theta)$$
$$M_n(\theta)=P(G_1'|\theta)+\cdots+P(G_n'|\theta)$$
$$N_n(\theta)=P(G_n|\theta)$$

および

(15.2.8)
$$L(\theta)=\lim_{n\to\infty}L_n(\theta)=P(G^\circ|\theta)$$
$$M(\theta)=\lim_{n\to\infty}M_n(\theta)=P(G'|\theta)$$
$$N(\theta)=\lim_{n\to\infty}N_n(\theta)$$

と置こう．ただし $G^\circ = G_1^\circ \cup G_2^\circ \cup \cdots$, $G' = G_1' \cup G_2' \cup \cdots$. $L_n(\theta)$, $M_n(\theta)$, $N_n(\theta)$ は非負で和が 1，しかも $L_n(\theta)$ と $M_n(\theta)$ は $n = 1, 2, \cdots$ に関して非減少だから，これらの極限は存在する．$L(\theta')$ および $M(\theta')$ は，母集団における θ の真値が θ' のとき，それぞれ $\mathscr{H}_0$ を採択する確率，および棄却する確率である．$L(\theta)$ は θ の関数と見れば，逐次検定 S の検査特性関数である．$N(\theta)$ は S に対する逐次過程が無限に続く確率を示している．$N(\theta) = 0$ は逐次過程が確率 1 で停止することを意味している．明らかに

(15.2.9)
$$L(\theta) + M(\theta) + N(\theta) = 1.$$

$G_n^\circ \cup G_n' = G_n^*$ とし

(15.2.10)
$$p(n|\theta) = P(G_n^*|\theta)$$

と置けば，$p(n|\theta)$ は，θ がパラメータの真値のとき，n 回の試行で S に対する逐次過程を（$\mathscr{H}_0$ を採択するか棄却するかの決定を行なって）停止させる確率になる．

S に対する逐次過程が確率 1 で停止する，すなわち $N(\theta) = 0$ ならば過程を停止するまでに要した試行の平均回数は

(15.2.11)
$$\mathscr{E}(n|\theta) = \sum_{t=1}^{\infty} t p(t|\theta)$$

になる．もちろんこれは θ の関数である．$\mathscr{E}(n|\theta)$ は逐次解析では通常，**平均検査個数**と呼ばれている．退化逐次検定，すなわち標本の大きさ n が固定されている検定の場合には，もちろん

$$\mathscr{E}(n|\theta) = n$$

である．$m = 1, 2, \cdots$ に対して

$$\mathscr{E}(n|\theta) \geqslant \sum_{t=1}^{m} t p(t|\theta) + mP(G_m|\theta)$$

だから $\mathscr{E}(n|\theta)$ が有限であるためには，$\lim_{m\to\infty} mP(G_m|\theta) = 0$，すなわち $N(\theta) = 0$ が必要である．

有限な平均 $\mathscr{E}(x)$ を持つ分布からの単純確率標本の場合については $\mathscr{E}(n)$ に関して Wald (1945) および Blackwell (1946) による次の定理がある．これは後の節でも用いる．

15.2.1　$(x_1, x_2, \cdots)$ は有限な平均 $\mathscr{E}(x)$ を持つ分布からの独立な確率変数列とし，S は $\mathscr{E}(n)$ が有限となる逐次検定であるとする．$x_1 + \cdots + x_n$ を S が停止するまで取り出された x の項の和とすれば

(15.2.12)
$$\mathscr{E}(x_1 + x_2 + \cdots + x_n) = \mathscr{E}(x)\cdot\mathscr{E}(n)$$

である．

15.2.1 の証明には

(15.2.13)
$$\mathscr{E}(x_1 + x_2 + \cdots + x_n) = \mathscr{E}(x_1) + \mathscr{E}(x_2|x_1 \in G_1)\cdot P(G_1) + \mathscr{E}(x_3|(x_1, x_2) \in G_2)\cdot P(G_2) + \cdots$$

に注意する．ただし $G_1, G_2, \cdots$ は (15.2.4) で定義されている．$x_1, x_2, \cdots$ は互いに独立で，すべて $\mathscr{E}(x)$ に等しい平均値を持つので (15.2.13) は

(15.2.14)
$$\mathscr{E}(x_1 + x_2 + \cdots + x_n) = \mathscr{E}(x)[1 + P(G_1) + P(G_2) + \cdots]$$

に帰着する．以前と同じく $G_n^* = G_n^\circ \cup G_n'$, $n = 1, 2, \cdots$ とする．このとき $G_1^*, \cdots, G_n^*$, G_n は R_∞ における排反な事象で，各 n に対するこれらの事象の確率の和は 1 で，しかも $\mathscr{E}(n)$ が有限だから，S は確率 1 で停止する．よって

(15.2.15)
$$1 = \sum_{t=1}^{\infty} P(G_t^*), \qquad P(G_t) = \sum_{t'=t+1}^{\infty} P(G_{t'}^*), \qquad t = 1, 2, \cdots.$$

これを (15.2.14) に代入すれば

(15.2.16)
$$\mathscr{E}(x_1 + x_2 + \cdots + x_n) = \mathscr{E}(x)\left[\sum_{t=1}^{\infty} tP(G_t^*)\right]$$

すなわち (15.2.12) が成立する．これで **15.2.1** の議論を終える．

(c)　逐次検定の選択に対する基準

逐次検定の形式的な定義においては G_n°, G_n', G_n が各 n について (15.2.1) を満足する排反な集合であるということを除けば，他に制限はなかった．逐次検定 S の構成に際して，基本的な問題の 1 つに G_n°, G_n' および G_n の選択がある．望ましくて，しかも常に満たされていなければならない要請の 1 つに，S に対する逐次過程が確率 1 で停止すること，すなわち

(15.2.17)
$$N(\theta) = 0$$

がある．

他に望ましい要請としては，すべての統計的検定の場合と同様に，第 1 種の誤りの危険を固定すること，すなわち

(15.2.18)
$$L(\theta_0) = 1 - \alpha$$

がある．

2つの逐次検定を比較する場合には（固定された大きさの標本に基づく統計的検定の場合と同じように）θ のある値，たとえば $\theta=\theta_1$ に対する第2種の誤りの危険 β を固定すること，すなわち

(15.2.19)　$$L(\theta_1)=\beta$$

を要請することも役立つだろう．θ_0 を他の θ_1 と対比して考えるとき，$\mathscr{H}_0$ は θ の真値が θ_0 であるという仮説を，また $\mathscr{H}_1$ は θ の真値が θ_1 であるという仮説を示すものとする．

条件 (15.2.17), (15.2.18) および (15.2.19) を仮定すれば，もちろん

(15.2.20)　$$M(\theta_0)=\alpha,\qquad M(\theta_1)=1-\beta$$

になる．

$\theta_0<\theta_1$ にとれば，それぞれ $\theta\leqslant\theta_0$, $\theta_0<\theta<\theta_1$, $\theta\geqslant\theta_1$ なる θ の値の集合は，$\mathscr{H}_0$ の採択領域，無関心領域および棄却領域と解釈できる．$\theta_0>\theta_1$ としても，同様な領域を設けることができる．

$N(\theta)=0$ を仮定したとき，条件 (15.2.17), (15.2.18), (15.2.19) の下で，$L(\theta)$ のグラフの基本的な型が図15.1に示されている．

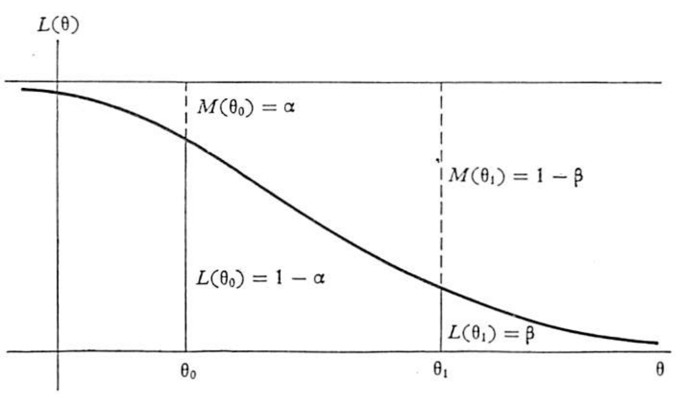

図 15.1

$L(\theta)$ が (15.2.18) と (15.2.19) を満たすような逐次検定を，強さ $(\alpha,\theta_0;\beta,\theta_1)$ の逐次検定と呼ぶ．

S_1, S_2 は検査特性関数 $L_1(\theta)$, $L_2(\theta)$ を持つ2つの逐次検定で，$L_1(\theta_0)=L_2(\theta_0)=1-\alpha$ で，一方 $L_1(\theta_1)<L_2(\theta_1)$ とする．このとき $\mathscr{H}_0$ を $\mathscr{H}_1$ に対して検定する S_1 は S_2 より強いという．すべての可能な検定 S_2 に対して，S_1 が S_2 より強いとき，S_1

（これを S^* で表わす）を最強力であるという．

強さが等しい2つの検定 S_1, S_2 の平均検査個数が $\mathscr{E}_1(n|\theta)$, $\mathscr{E}_2(n|\theta)$ で

(15.2.21)　$$\mathscr{E}_1(n|\theta)\leqslant\mathscr{E}_2(n|\theta),\qquad \theta=\theta_0,\theta_1$$

（ただし等号は θ_0, θ_1 の両方に対して同時には成り立たない）ならば，S_1 は $\mathscr{H}_0$ を $\mathscr{H}_1$ に対して検定するとき，S_2 より好ましいと考えられる．

θ_0, θ_1 の両方と，S_1 と異なるすべての可能な S_2 に対して関係式 (15.2.21) が成り立つならば，S_1（これを S^* で表わす）は最大効果の逐次検定である．15.4節では，このような最大効果の逐次検定 S^* が確率比に基づく方法によって構成されることを示そう．

15.3　カルテシアン逐次検定

(a)　検定の表現および性質

考えられる最良の可能な逐次検定を取り扱う前に，たとえ効果的でないにしても2項待ち時間分布 (6.5.15) に基づく簡単な逐次検定を議論することは考え方を定着させる意味で有益であるだろう．$(x_1,x_2,\cdots)$ はすべて同じ c.d.f. $F(x;\theta)$ を持つ独立な確率変数列であるとする．

$\mathscr{H}_0$ を $\mathscr{H}_1$ に対して検定する次の逐次検定 S を考えよう．G_0°, G_0', G_0 を和集合が全実数軸であるような排反な初期集合とし，各 $n=1,2,\cdots$ について $G_{(n)}^\circ$, $G_{(n)}'$, $G_{(n)}$ をそれぞれ $R_1^{(n)}$ における集合 G_0°, G_0', G_0 とする．ただし各 n について $R_1^{(n)}$ は x_n の標本空間である．このとき (15.2.4) で定義された排反な事象 $G_1^\circ,\cdots,G_n^\circ$, $G_1',\cdots,G_n'$, G_n が次のように構成できる．

(15.3.1)
$$\begin{aligned}
G_1^i &= \{x_1\in G_{(1)}^i\}, \qquad i={}^\circ,{}' \\
G_2^i &= \{(x_1,x_2)\in G_{(1)}\times G_{(2)}^i\}, \qquad i={}^\circ,{}' \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
G_n^i &= \{(x_1,\cdots,x_n)\in G_{(1)}\times\cdots\times G_{(n-1)}\times G_{(n)}^i\}, \qquad i={}^\circ,{}' \\
G_n &= \{(x_1,\cdots,x_n)\in G_{(1)}\times\cdots\times G_{(n)}\}.
\end{aligned}$$

したがって S に対する逐次過程は $(x_1,x_2,\cdots)$ から x を1つ取り出したとき，その x が実数軸上の G_0 にはいる限り，次の x を取り続けることになる．x が G_0° にはいれば $\mathscr{H}_0$ を採択し，一方 x が G_0' にはいれば $\mathscr{H}_1$ を採択する．このような検定をカルテシアン逐次検定と呼ぶ．さて

$$a(\theta)=\int_{G_0^\circ}dF(x;\theta),\qquad b(\theta)=\int_{G_0'}dF(x;\theta) \tag{15.3.2}$$

$$c(\theta)=\int_{G_0}dF(x;\theta)=1-a(\theta)-b(\theta)$$

としよう．$a(\theta)$, $b(\theta)$, $c(\theta)$ を a, b および c で表わし，$a(\theta_i)=a_i$, $b(\theta_i)=b_i$, $c(\theta_i)=c_i$, $i=0,1$ とする．こうすれば

$$P(G_t^\circ|\theta)=ac^{t-1},\qquad P(G_t'|\theta)=bc^{t-1},\qquad P(G_n|\theta)=c^n \tag{15.3.3}$$

$$t=1,\cdots,n$$

$$L_n(\theta)=\frac{a(1-c^n)}{1-c},\qquad M_n(\theta)=\frac{b(1-c^n)}{1-c},\qquad N_n(\theta)=c^n. \tag{15.3.4}$$

G_0 を $c(\theta)<1$ となるように選べば，$\lim_{n\to\infty}N_n(\theta)=0$ となり

$$L(\theta)=\frac{a}{a+b},\qquad M(\theta)=\frac{b}{a+b} \tag{15.3.5}$$

である．

x_n を取り出したことによって逐次過程が停止する確率は

$$p(n|\theta)=(a+b)c^{n-1} \tag{15.3.6}$$

で与えられる．そして平均検査個数は

$$\mathscr{E}(n|\theta)=(a+b)\sum_{n=1}^{\infty}nc^{n-1}=\frac{1}{a+b} \tag{15.3.7}$$

になる．$p(n|\theta)$ は 2 項待ち時間分布 (6.5.15) で $k=1$ とした特別の場合に過ぎない．

第 1 種の誤りの危険，および第 2 種の誤りの危険がそれぞれ $\theta=\theta_0$ で α, $\theta=\theta_1$ で β ならば，G_0°, G_0' は方程式

$$\frac{a_0}{a_0+b_0}=1-\alpha,\qquad \frac{a_1}{a_1+b_1}=\beta \tag{15.3.8}$$

を満足しなければならない．すなわち

$$\frac{a_0}{b_0}=\frac{1-\alpha}{\alpha},\qquad \frac{a_1}{b_1}=\frac{\beta}{1-\beta} \tag{15.3.9}$$

になる．

α, β および $\mathscr{E}(n|\theta_0)$ $(=1/(a_0+b_0))$ の個々の定められた値に対して，G_0°, G_0' の選択全体が存在すれば，$F(x;\theta)$ が p.d.f. $f(x;\theta)$ を持つとき，最強力カルテシアン逐次検定を与える選択方法がネイマン=ピヤソンの定理 **13.2.1** で示される．すなわち G_0° として確率比が

$$\frac{f(x;\theta_1)}{f(x;\theta_0)}\leqslant k_0 \tag{15.3.10}$$

を満たす x の値の集合をとり，G_0' として

$$\frac{f(x;\theta_1)}{f(x;\theta_0)}\geqslant k_1 \tag{15.3.11}$$

を満たす場合をとる．ただし $k_1>k_0>0$ は α, β, $\mathscr{E}(n|\theta_0)$ の定められた値に対して G_0° と G_0' が (15.3.8) を満たすように選ばれる．いい換えれば，$a_0/(a_0+b_0)$ を α に，$\mathscr{E}(n|\theta_0)$ を n_0 に固定する（したがって第 1 種の誤りの確率と θ_0 における平均検査個数を固定する）ような集合 G_0°, G_0' の任意の他の選択 $G_0^{\circ *}$, G_0^{1*} は $a_1/(a_1+b_1)>\beta$ なるカルテシアン逐次検定 S を与える．すなわちこの場合の第 2 種の誤りの確率は (15.3.10), (15.3.11) で与えられた G_0°, G_0' に基づく逐次検定の第2種の誤りの確率を超える．

この事柄は本質的には **13.2.1** と同じ仮定の下で成り立つ．証明も **13.2.1** と同様である．読者に練習問題として残しておこう [問題 15.6]．

注意　母集団から一度に r 個の x を取り出すことによって r 重カルテシアン逐次検定 S を考えることができる．$(x_1,\cdots,x_r)$ の標本空間 R_r における採択集合および棄却集合が確率比で定められる場合だけを考えれば十分であろう．さて G_0° を

$$\prod_{i=1}^{r}\frac{f(x_i;\theta_1)}{f(x_i;\theta_0)}\leqslant k_0 \tag{15.3.12}$$

を満たす $(x_1,\cdots,x_r)$ の集合，G_0' を

$$\prod_{i=1}^{r}\frac{f(x_i;\theta_1)}{f(x_i;\theta_0)}\geqslant k_1 \tag{15.3.13}$$

を満たす集合とする．ただし G_0 はどちらの不等式も満たさない集合である．k_1, k_0 は $P(G_0^\circ|\theta)$, $P(G_0'|\theta)$, $P(G_0|\theta)$ でそれぞれ $a(\theta)$, $b(\theta)$, $c(\theta)$ を定義して，(15.3.8) を満足するように $k_1>k_0>0$ を選ぶ．このように G_0°, G_0' を選べば，**13.2.1** と同様な議論によって $(x_1,\cdots,x_r)$ の標本空間において，$a_0/(a_0+b_0)$ を α に固定し $\mathscr{E}(n|\theta_0)$ も固定するような，G_0°, G_0' の他のどの選択方法よりも強いカルテシアン逐次検定が得られることがわかる．この検定は $s<k$ なる任意の s 重カルテシアン逐次検定よりも強い．

読者には次の事柄は明らかだろう．1重カルテシアン逐次検定に対して第1種の誤り，第2種の誤りを α, β に固定することが困難ならば，r を適当に選んで，r 重カルテシアン逐次検定を用いれば，この 2 つの誤りをいくらでも α, β に近づけることができる．

(b)　ノンパラメトリック検定への応用

この節の冒頭で述べたように，カルテシアン逐次検定は考えられ得る最良の逐次検定ではないかもしれない．しかし $F(x;\theta)$ の連続性以外に何もわかっていないならば，カルテシアン逐次検定はノンパラメトリック逐次検定を与えることになる．これを見るために $F(x;\theta_0)$，$F(x;\theta_1)$ を連続とし，それぞれ $F_0(x)$，$F_1(x)$ で表わそう．G_0°，G_0' を区間 $(-\infty,a)$，$(b,+\infty)$，$a<b$ に選ぶと (15.3.8) は

(15.3.14)
$$\frac{F_0(a)}{F_0(a)+1-F_0(b)}=1-\alpha,\qquad \frac{F_1(a)}{F_1(a)+1-F_1(b)}=\beta$$

と書かれる．G_0°，G_0' のこの選択に基づくカルテシアン逐次検定は対立仮説《a，b でそれぞれ値 $F_1(a)$，$1-(1-\beta)F_1(a)/\beta$ を持つ連続な c.d.f. $F_1(x)$ を採択する》に対して仮説《a，b でそれぞれ値 $F_0(a)$，$1-\alpha F_0(a)/(1-\alpha)$ を持つ c.d.f. $F_0(x)$ を採択する》の検定になる．平均検査個数は $F_0(x)$ が実際抽出された分布ならば $[F_0(a)+1-F_0(b)]^{-1}$，$F_1(x)$ ならば $[F_1(a)+1-F_1(b)]^{-1}$ である．

15.4　逐次確率比検定

(a)　検定の定義

前節では，G_0°，G_0'，G_0 を選んで $\mathscr{H}_1$ に対して $\mathscr{H}_0$ を検定する最強力なカルテシアン逐次検定を導くような方法を議論した．これは逐次検定に対して集合 $G_1^\circ,\cdots,G_n^\circ$，$G_1',\cdots,G_n'$，$G_n$，$n=1,2,\cdots$ の最良の選択方法はどういうものであろうかということを示している．事実 G_n°，G_n'，G_n として，各 n に対し次のおのおのの不等式で定義された(15.2.1)を満たす R_n の排反な事象を選べばよい．

(15.4.1)
$$\frac{Q_{1n}}{Q_{0n}}\leqslant k_0,\qquad \frac{Q_{1n}}{Q_{0n}}\geqslant k_1,\qquad k_0<\frac{Q_{1n}}{Q_{0n}}<k_1.$$

ただし

(15.4.2)
$$Q_{in}=\prod_{t=1}^{n}f(x_t;\theta_i),\qquad i=0,1$$

で $f(x;\theta_0)$，$f(x;\theta_1)$ はともに p.f. であるか，または絶対連続な p.d.f. である．この節では p.d.f. の場合だけ考えれば十分である．p.d.f. に対する結果がそのまま p.f. の場合にも当てはまるからである．

Wald (1945) によってはじめて提唱されたこの検定は逐次確率比検定と呼ばれている．検定方式は $(x_1,\cdots,x_n)$ $(n=1,2,\cdots)$ が x_n を取り出したとき，3 つのうちのどの不等式を満たすかに応じて $\mathscr{H}_0$ を採択したり，$\mathscr{H}_1$ を採択したり，あるいは x_{n+1} を抜き取るものである．ただし k_0，k_1 は

(15.4.3)
$$L(\theta_0)=1-\alpha,\qquad L(\theta_1)=\beta$$

となるように固定する．すなわち第 1 種の誤りの確率，第 2 種の誤りの確率が α，β になるようにする．

(b)　逐次確率比過程を停止するまでの検査個数の分布関数の性質

まず逐次確率比過程が確率 1 で停止するかどうかという問題を考えよう．

(15.4.4)
$$z=\log\left(\frac{f(x;\theta_1)}{f(x;\theta_0)}\right)$$

とし，$H(z;\theta)$ をパラメータ空間 Ω において値を θ に固定したときの z の c.d.f. とする．$\theta_0\neq\theta_1$ のとき z は非退化確率変数であるとする．

(15.4.5)
$$z_t=\log\left(\frac{f(x_t;\theta_1)}{f(x_t;\theta_0)}\right),\qquad t=1,2,\cdots$$

とすれば $(z_1,z_2,\cdots)$ は c.d.f. $H(z;\theta)$ からの単純確率標本によって生成される確率過程である．この逐次過程は不等式

(15.4.6)
$$\log k_0<z_1+\cdots+z_n<\log k_1$$

が満たされなくなる最小の n のとき停止する．G_n は不等式 (15.4.6) が成り立つ事象だから，G_n は逐次過程が停止するまでに n 個よりも多く x を抜き取る事象になる．

ある整数 r に対し，$n=mr$ とし，$\zeta_1=z_1+\cdots+z_r$，$\zeta_2=z_{r+1}+\cdots+z_{2r}$，$\cdots$，$\zeta_m=z_{(m-1)r+1}+\cdots+z_{mr}$，$\cdots$ とする．このとき，I_{mr} を

(15.4.7)
$$|\zeta_i|<D,\qquad i=1,2,\cdots,m$$

なる事象とする．ただし $D=\max(|\log k_0|,|\log k_1|)$．また $(\zeta_1,\zeta_2,\cdots,\zeta_m)$ はある c.d.f. $J(\zeta;\theta)$ からの標本になる．$G_{mr}\subset I_{mr}$ だから，

(15.4.8)
$$N_{mr}(\theta)=P(G_{mr}|\theta)\leqslant P(I_{mr}|\theta)=p^m.$$

ただし，$p=P(|\zeta|<D)$．

z は非退化だから正の分散 σ^2 を持つ．ζ の分散は $r\sigma^2$．これは r を $r>D^2/\sigma^2$ となるように選べば D^2 を超える．$\mathscr{E}(\zeta^2)>D^2$．よって

$$P(\zeta^2<D^2)=p<1.$$

ゆえに $r > D^2/\sigma^2$ として $m \longrightarrow \infty$ とすれば，逐次過程が無限に続く確率は $\lim_{m\to\infty} p^m = 0$ になる．

要約すると，

15.4.1　(15.4.4) で定義された確率変数 z が非退化（すなわち，分散が正）ならば，逐次確率比過程は確率 1 で停止する．

事実，逐次過程を停止させるまでに要する試行の回数の分布関数について，これよりも強い Stein (1946) による定理が成立している．スタインの定理の変形は次のようになる．

15.4.2　確率変数 z が非退化ならば，n のすべてのモーメントが存在する．

15.4.2 を示すには n のモーメント母関数

(15.4.9)
$$\phi(u) = \mathscr{E}(e^{un}) = \sum_{t=1}^{\infty} e^{ut} P(G_t^* | \theta)$$

を考えれば十分である．ただし $G_t^* = G_t^\circ \cup G_t'$ は t 番目の抜き取りで逐次過程が停止する事象である．(15.4.9) における右辺の級数の r 個の項をひとまとめにして評価すれば，$u \geqslant 0$ に対して

$$\phi(u) \leqslant e^{ru}P(0 < n \leqslant r) + e^{2ru}P(r < n \leqslant 2r) + \cdots$$
$$\leqslant e^{ru}P(n > 0) + e^{2ru}P(n > r) + \cdots$$

となる．しかし (15.4.8) より，z が正の分散を持てば

$$P(n \geqslant mr) \leqslant p^m.$$

したがって

(15.4.10)
$$\phi(u) \leqslant e^{ru} + e^{2ru}p + e^{3ru}p^2 + \cdots = e^{ru}(1 - e^{ru}p)^{-1}.$$

$u \leqslant 0$ のときも同様に

(15.4.10 *a*)
$$\phi(u) \leqslant e^{u}(1 - e^{ru}p)^{-1}.$$

ゆえに u を $e^{ru}p < 1$ を満たす任意の（実）数とすれば，$\phi(u)$ は収束する．不等式 $e^{ru}p < 1$ が成立する u の区間は $u = 0$ を内点として含み，$u = 0$ における $\phi(u)$ の k 次微係数が存在し，k 次のモーメントを与える．ただし $k = 1, 2, \cdots$.

最後に（**15.2.1** の系としても述べられる）次の結果は，$\mathscr{E}(n|\theta)$ を求める場合や逐次確率比検定の有効性を吟味する場合に重要な役割を果たす．

15.4.3　z が非退化確率変数ならば

(15.4.11)
$$\mathscr{E}(z_1 + z_2 + \cdots + z_n | \theta) = \mathscr{E}(z|\theta) \cdot \mathscr{E}(n|\theta)$$

である．

$\mathscr{E}(n|\theta)$ の有限性は **15.4.2** からただちにわかる．あとの議論は **15.2.1** により成立する．

(c)　逐次確率比検定に対する境界定数 k_0, k_1 の決定

さて強さ $(\alpha, \theta_0 ; \beta, \theta_1)$ の逐次確率比検定に対する境界定数 k_0 と k_1 の値を決定する問題を考えよう．k_0 と k_1 を正確に定めることは難しい問題である．しかし近似的な不等式と精度の高い近似は容易に見い出せる．

逐次確率比検定の定義より，(15.4.1) の最初の不等式が満たされるならば，$\mathscr{H}_0$ は x_n の抜き取りに基づいて採択されるだろう．この事象を G_n° で表わしているので，G_n° のすべての点で

(15.4.12)
$$Q_{1n} \leqslant k_0 Q_{0n}, \qquad n = 1, 2, \cdots$$

が成り立つ．(15.2.8) より

(15.4.13)
$$L(\theta) = P(G_1^\circ | \theta) + P(G_2^\circ | \theta) + \cdots.$$

$L(\theta)$ の強さが $(\alpha, \theta_0 ; \beta, \theta_1)$ ならば

(15.4.14)
$$L(\theta_0) = 1 - \alpha, \qquad L(\theta_1) = \beta$$

でなければならない．(15.4.12) から

(15.4.15)
$$P(G_n^\circ | \theta_1) \leqslant k_0 P(G_n^\circ | \theta_0), \qquad n = 1, 2, \cdots$$

が成立する．ゆえに

$$\sum_{n=1}^{\infty} P(G_n^\circ | \theta_1) \leqslant k_0 \sum_{n=1}^{\infty} P(G_n^\circ | \theta_0), \qquad n = 1, 2, \cdots$$

すなわち

(15.4.16)
$$L(\theta_1) \leqslant k_0 L(\theta_0).$$

(15.4.14) を代入すれば，k_0 は不等式

$$k_0 \geqslant \frac{\beta}{1 - \alpha}$$

を満足しなければならない．

同様な方法を用いて，x_n $(n = 1, 2, \cdots)$ の抜き取りに応じて $\mathscr{H}_1$ を採択する事象 G_n' を考察する．すなわち (15.4.1) の第 2 の不等式が成立するとき，$\mathscr{H}_1$ を採択すれば

(15.4.17)
$$M(\theta_1) \geqslant k_1 M(\theta_0)$$
が得られる．$N(\theta)=0$ より，$M(\theta)=1-L(\theta)$．これを (15.4.17) に代入して (15.4.14) を用いれば，k_1 は
$$k_1 \leqslant \frac{1-\beta}{\alpha}$$
を満たさねばならない．したがって

15.4.4　強さ $(\alpha, \theta_0 ; \beta, \theta_1)$ の逐次確率比検定に対する定数 k_0 と k_1 は，不等式

(15.4.18)
$$k_0 \geqslant \frac{\beta}{1-\alpha}, \qquad k_1 \leqslant \frac{1-\beta}{\alpha}$$
を満足する．

さて α と β が与えられたとき，逐次確率比検定の実際の強さを調べよう．k_0，k_1 を

(15.4.19)
$$k_0 = \frac{\beta}{1-\alpha}, \qquad k_1 = \frac{1-\beta}{\alpha}$$
に選ぶ．便宜上，$(\alpha, \theta_0 ; \beta, \theta_1)$ を検定の意図された強さ，$(\alpha', \theta_0 ; \beta', \theta_1)$ を検定の実際の強さと考えよう．このとき (15.4.18) より

(15.4.20)
$$\frac{\beta}{1-\alpha} \geqslant \frac{\beta'}{1-\alpha'}, \qquad \frac{1-\beta}{\alpha} \leqslant \frac{1-\beta'}{\alpha'}$$
である．この 2 つの不等式より

(15.4.21)
$$\beta' \leqslant \frac{\beta}{1-\alpha}, \qquad \alpha' \leqslant \frac{\alpha}{1-\beta}$$
で

(15.4.22)
$$\alpha' + \beta' \leqslant \alpha + \beta$$
も成立する．最後の不等式は (15.4.19) で与えられた k_0，k_1 を選べば実際の第 1 種の誤りの確率と第 2 種の誤りの確率の和が意図された第 1 種の誤りの確率と第 2 種の誤りの確率の和を超えないことを意味している．したがって次を得る．

15.4.5　意図された強さ $(\alpha, \theta_0 ; \beta, \theta_1)$ の逐次確率比検定に対する定数 k_0，k_1 を実際にそれぞれ $\beta/(1-\alpha)$，$\alpha/(1-\beta)$ として選べば，この検定の実際の強さは $(\alpha', \theta_0 ; \beta', \theta_1)$ である．ただし α'，β' は不等式 (15.4.21)，(15.4.22) を満たす．

実際の応用例では，意図された強さ $(\alpha, \theta_0 ; \beta, \theta_1)$ の α および β の値はほとんど 0.10 を超えることはない．また実際の強さと意図された強さとの違いは実用的な見地に立てば，大して重要でないことを不等式 (15.4.21)，(15.4.22) が示している．

(d)　逐次確率比検定の検査特性関数

これまでは検査特性関数 $L(\theta)$ を θ の 2 つの値 θ_0，θ_1 だけに対してしか考えなかった．パラメータ空間 Ω の任意の θ について $L(\theta)$ の値を完全に定めることは困難な問題である．しかし $L(\theta)$ の近似表現は比較的容易に求めることができる．

はじめに Wald (1945) による次の補題を示しておくと便利であろう．

15.4.6　$z=\log f(x ; \theta_1) - \log f(x ; \theta_0)$ とする．ただし x は p.d.f. $f(x, \theta)$ を持つ確率変数，θ は θ_0，θ_1 を含むパラメータ空間 Ω の点とする．

(i)　$\mathscr{E}(e^{zh})$ がすべての実数 h に対して存在し

(ii)　$\mathscr{E}(z) \neq 0$

(iii)　$\delta_1 > 0$，$0 < \delta_2 < 1$ を満たすある δ_1，δ_2 に対して $P(e^z > 1+\delta_1) > 0$ かつ $P(e^z < 1-\delta_2) > 0$

が成立すれば，各 θ に対し h，たとえば $h(\theta) \neq 0$ が存在して

(15.4.23)
$$\int_{-\infty}^{\infty} \left(\frac{f(x ; \theta_1)}{f(x ; \theta_0)}\right)^h f(x ; \theta)\, dx = 1$$
すなわち
$$\mathscr{E}(e^{zh(\theta)}) = 1$$
となる．

15.4.6 を証明しよう．E_1 を $e^z > 1+\delta_1$ なる x の値の集合，E_2 を $e^z < 1-\delta_2$ を満たす x の値の集合とする．このとき $h>0$ に対して

(15.4.24)
$$\mathscr{E}(e^{zh}) \geqslant \int_{E_1} e^{zh} f(x ; \theta)\, dx > (1+\delta_1)^h P(E_1).$$
$h<0$ に対して

(15.4.25)
$$\mathscr{E}(e^{zh}) \geqslant \int_{E_2} e^{zh} f(x ; \theta)\, dx > (1-\delta_2)^h P(E_2).$$
$P(E_1)$ と $P(E_2)$ は 0 でないから，(15.4.24) および (15.4.25) より

(15.4.26)
$$\lim_{h\to+\infty} \mathscr{E}(e^{zh}) = \lim_{h\to-\infty} \mathscr{E}(e^{zh}) = +\infty$$

は明らか．$\mathscr{E}(e^{zh})$ を $\phi(h)$，$\phi(h)$ の1次および2次導関数を $\phi'(h)$，$\phi''(h)$ とすれば

(15.4.27) $$\phi'(0) = \mathscr{E}(z) \neq 0$$

(15.4.28) $$\phi''(h) = \mathscr{E}(z^2 e^{zh}) > 0$$

である．$\phi(0) = 1$，$\phi'(0) \neq 0$ かつ $\phi''(h) > 0$ だから，明らかに $\phi(h)$ は h ばかりでなく θ にも依存するが，1 より小さい最小値を持つ．さらに $\phi(h) - 1 = 0$ は 2 根 $h = 0$ と $h = h(\theta)$（ただし $h(\theta) \neq 0$）を持つこともわかる．これで **15.4.6** の議論を終わる．

さて $L(\theta)$ を近似する問題を考察しよう．x_n の抜き取りによって $\mathscr{H}_0$ が採択される事象 G_n° は (15.4.1) の最初の不等式で定められる．同じ事象 G_n° は，不等式

(15.4.29) $$Q_{*n} \leqslant k_0^h Q_n$$

で定められる．ただし

(15.4.30) $$Q_{*n} = \left(\frac{Q_{1n}}{Q_{0n}}\right)^h Q_n, \qquad Q_n = \prod_{t=1}^{n} f(x_t\,;\,\theta)$$

で，h は任意の実数である．この Q_{*n} は次のように書ける．

(15.4.31) $$Q_{*n} = \prod_{t=1}^{n}\left[\left(\frac{f(x_t\,;\,\theta_1)}{f(x_t\,;\,\theta_0)}\right)^h f(x_t\,;\,\theta)\right].$$

15.4.6 の条件の下では各 θ に対し h，たとえば $h(\theta) \neq 0$ が存在して

(15.4.32) $$\int \left(\frac{f(x\,;\,\theta_1)}{f(x\,;\,\theta_0)}\right)^{h(\theta)} f(x\,;\,\theta)\,dx = 1$$

となる．したがって

(15.4.33) $$f^*(x\,;\,\theta) = \left(\frac{f(x\,;\,\theta_1)}{f(x\,;\,\theta_0)}\right)^{h(\theta)} f(x\,;\,\theta)$$

で定義される関数 $f^*(x\,;\,\theta)$ は p.d.f. の性質を持つ．

$P(G_n^\circ|\theta)$ が $f(x\,;\,\theta)$ で評価された事象 G_n° の確率であったように $P^*(G_n^\circ|\theta)$ を $f^*(x\,;\,\theta)$ から計算した G_n° の確率とする．(15.2.8) の $L(\theta)$ と同様に次式で $L^*(\theta)$ を定義しよう．

$$L^*(\theta) = \sum_{n=1}^{\infty} P^*(G_n^\circ|\theta) = P^*(G^\circ|\theta).$$

そのとき (15.4.16) を導いた方法と同様に議論すれば

(15.4.34) $$L^*(\theta) \leqslant k_0^{h(\theta)} L(\theta)$$

が成り立つ．

事象 G_n' は $\mathscr{H}_1$ を採択する結果になり，(15.4.1) の 2 番目の式で定義される．この事象は不等式

(15.4.35) $$Q_{*n} \geqslant k_1^h Q_n$$

で定義しても同値である．(15.4.34) を得た方法と同様な方針で論ずれば，(15.4.17) に対応して

(15.4.36) $$1 - L^*(\theta) \geqslant k_1^{h(\theta)}(1 - L(\theta))$$

が得られる．

k_0，k_1 の近似値を求めるのに，(15.4.18) の不等式を (15.4.19) の等式に置き換えたことから，同様に類推すれば，$L(\theta)$ を近似するために不等式 (15.4.34)，(15.4.36) を等式に置き換えればよかろう．この近似は

(15.4.37) $$L(\theta) \cong \frac{1 - k_1^{h(\theta)}}{k_0^{h(\theta)} - k_1^{h(\theta)}}$$

となることがわかる．

$h(\theta_0) = 1$，$h(\theta_1) = -1$ だから，k_0，k_1 の値を (15.4.19) で近似すれば，近似式 (15.4.37) は強さ $(\alpha, \theta_0\,;\,\beta, \theta_1)$ の逐次確率比検定に対する $L(\theta_0)$，$L(\theta_1)$ の正しい値，すなわち (15.4.14) の値を与える．

$L(\theta)$ の近似式 (15.4.37) は実用上の目的に対しては満足できるものである．θ_0，θ_1 と異なる θ の値に対するこの近似式の精度を正確に調べることは莫大な解析を必要とするためここでは省略する．これに関する詳細な解析は Wald (1947 a) の著作に見られる．

要約すると次のようになる．

15.4.7　**15.4.6** の条件下では，強さ $(\alpha, \theta_0\,;\,\beta, \theta_1)$ の逐次確率比検定に対する検査特性関数 $L(\theta)$ の近似式は (15.4.37) で与えられる．ただし $h(\theta)$ は (15.4.32) を満足し，k_0，k_1 は (15.4.19) で近似されている．

(e)　逐次確率比検定の平均検査個数

$\mathscr{E}(n|\theta)$ の近似表現を得るために次のように議論を進めよう．$\mathscr{E}(z|\theta) \neq 0$ ならば (15.4.11) より

(15.4.38) $$\mathscr{E}(n|\theta) = \frac{\mathscr{E}(z_1 + z_2 + \cdots + z_n|\theta)}{\mathscr{E}(z|\theta)}.$$

以前と同様に $G^\circ = G_1^\circ \cup G_2^\circ \cup \cdots$，$G' = G_1' \cup G_2' \cup \cdots$ とする．このとき $P(G^\circ|\theta) = L(\theta)$，$P(G'|\theta) = 1 - L(\theta)$ である．また

(15.4.39) $$\mathscr{E}(z_1 + z_2 + \cdots + z_n|\theta) = \mathscr{E}(z_1 + z_2 + \cdots + z_n|G^\circ\,;\,\theta)L(\theta)$$

$$+\mathscr{E}(z_1+z_2+\cdots+z_n|G';\theta)(1-L(\theta)).$$

しかし

$$\mathscr{E}(z_1+z_2+\cdots+z_n|G^\circ;\theta)\leqslant \log k_0$$

かつ

$$\mathscr{E}(z_1+z_2+\cdots+z_n|G';\theta)\geqslant \log k_1.$$

この不等式の k_0, k_1 を (15.4.19) の値で置き換え，不等式を等式にすると (15.4.38) より $\mathscr{E}(n|\theta)$ に対する次の近似式が得られる．

(15.4.40)
$$\mathscr{E}(n|\theta)\cong\frac{L(\theta)\log\left(\frac{\beta}{1-\alpha}\right)+(1-L(\theta))\log\left(\frac{1-\beta}{\alpha}\right)}{\mathscr{E}(z|\theta)}.$$

θ_0, θ_1 における $\mathscr{E}(n|\theta)$ の近似値は

(15.4.41)
$$\mathscr{E}(n|\theta_0)\cong\frac{(1-\alpha)\log\left(\frac{\beta}{1-\alpha}\right)+\alpha\log\left(\frac{1-\beta}{\alpha}\right)}{\mathscr{E}(z|\theta_0)}$$
$$\mathscr{E}(n|\theta_1)\cong\frac{\beta\log\left(\frac{\beta}{1-\alpha}\right)+(1-\beta)\log\left(\frac{1-\beta}{\alpha}\right)}{\mathscr{E}(z|\theta_1)}$$

である．

この近似は実用的な目的に対しては満足すべきものである．(15.4.40) の近似式の誤差の限界を求める問題は相当な解析を必要とするためここでは省略する．この詳細に興味を持たれる読者には Wald (1947 a) が参考になるだろう．

(f) 逐次確率比検定の有効性

Wald (1945) による次の定理は任意の逐次検定に対して，$\mathscr{E}(n|\theta)$ の θ_0 および θ_1 における下界を与えている．

15.4.8　$(x_1,x_2,\cdots)$ は $\mathscr{E}(z|\theta)\neq 0$ なる p.d.f. $f(x;\theta)$ を持つ独立な確率変数列とする．ただし z は (15.4.4) で定義されている．S を確率 1 で停止する強さ $(\alpha,\theta_0;\beta,\theta_1)$ の任意の逐次検定とすれば

(15.4.42)
$$\mathscr{E}(n|\theta_0)\geqslant\frac{(1-\alpha)\log\left(\frac{\beta}{1-\alpha}\right)+\alpha\log\left(\frac{1-\beta}{\alpha}\right)}{\mathscr{E}(z|\theta_0)}$$
$$\mathscr{E}(n|\theta_1)\geqslant\frac{\beta\log\left(\frac{\beta}{1-\alpha}\right)+(1-\beta)\log\left(\frac{1-\beta}{\alpha}\right)}{\mathscr{E}(z|\theta_1)}$$

が成立する．

$(x_1,\cdots,x_n)$ の標本空間において (15.2.1) を満たす基本事象を G_n°, G_n', G_n, $n=1,2,\cdots$ で表わしても混乱は起らないだろう．このような G の項の列によって勝手な逐次検定 S が定義される．[S が逐次確率比検定の場合には G_n°, G_n', G_n, $n=1,2,\cdots$ は (15.2.1) の等式と (15.4.1) の 3 つの不等式で定義される.]

いつものように $z_1,z_2,\cdots$ を (15.4.5) で定義する．(15.4.11) より S に対して

(15.4.43)
$$\mathscr{E}(n|\theta)=\frac{\mathscr{E}(z_1+z_2+\cdots+z_n|\theta)}{\mathscr{E}(z|\theta)}.$$

しかし

(15.4.44)
$$\mathscr{E}(z_1+z_2+\cdots+z_n|\theta)=\mathscr{E}(z_1+z_2+\cdots+z_n|G^\circ;\theta)P(G^\circ|\theta)+\mathscr{E}(z_1+z_2+\cdots+z_n|G';\theta)P(G'|\theta).$$

まず $\mathscr{E}(n|\theta_0)$ を考えよう．S の強さは $(\alpha,\theta_0;\beta,\theta_1)$ だから

(15.4.45)
$$P(G^\circ|\theta_0)=1-\alpha,\qquad P(G'|\theta_0)=\alpha.$$

任意の確率変数 y について

(15.4.46)
$$\mathscr{E}(y)\leqslant\log\mathscr{E}(e^y)$$

が成り立つことを利用すれば

(15.4.47)
$$\mathscr{E}(z_1+z_2+\cdots+z_n|G^\circ;\theta_0)\leqslant\log\mathscr{E}(e^{z_1+z_2+\cdots+z_n}|G^\circ;\theta_0)=\log\mathscr{E}\left(\frac{Q_{1n}}{Q_{0n}}\middle|G^\circ;\theta_0\right).$$

ただし $\mathscr{E}\left(\frac{Q_{1n}}{G_{0n}}\middle|G^\circ;\theta_0\right)$ は $\theta=\theta_0$ のとき逐次過程が停止するまでの積

$$\left(\frac{f(x_1;\theta_1)}{f(x_1;\theta_0)}\right)\left(\frac{f(x_2;\theta_1)}{f(x_2;\theta_0)}\right)\cdots$$

の $G^\circ(=G_1^\circ\cup G_2^\circ\cup\cdots)$ 上の条件つき平均値である．しかし

(15.4.48)
$$\mathscr{E}\left(\frac{Q_{1n}}{Q_{0n}}\middle|G^\circ;\theta_0\right)=P(G^\circ|\theta_1)/P(G^\circ|\theta_0)=\frac{\beta}{1-\alpha}.$$

ゆえに

(15.4.49)
$$\mathscr{E}(z_1+z_2+\cdots+z_n|G^\circ;\theta_0)\leqslant\log\frac{\beta}{1-\alpha}.$$

同様に

(15.4.50)
$$\mathscr{E}(z_1+z_2+\cdots+z_n|G';\theta_0)\leqslant\log\frac{1-\beta}{\alpha}.$$

(15.4.45), (15.4.49), (15.4.50) を (15.4.44) に代入すれば (15.4.42) の最初の不等

式を得る．2 番目の不等式も同様にすれば得られる．これで **15.4.8** の議論を終える．

$\mathscr{E}(n|\theta)$ の下界に関しては，さらに Hoeffding (1960), Kiefer と Weiss (1957) によって研究されている．Anderson (1960) と Donnelly (1957) は分散既知，平均未知の正規分布からの標本抽出の場合における確率検定を変形して $\mathscr{E}(n|\theta)$ に関する問題に帰着させている．

逐次確率比検定，すなわち G_n°, G_n', G_n が (15.2.1) および (15.4.1) の不等式で定められている場合には

(15.4.51)
$$\mathscr{E}(z_1+z_2+\cdots+z_n|G^\circ;\theta)\leqslant\log k_0$$
$$\mathscr{E}(z_1+z_2+\cdots+z_n|G';\theta)\geqslant\log k_1$$

が成り立つことを見た．

x_n の抜き取りに基づいて標本点 $(x_1,\cdots,x_n)$ が G_n° に落ちるとき（すなわち $z_1+\cdots+z_n\leqslant\log k_0$ のとき），$z_1+\cdots+z_n$ に値 $\log k_0$ を与え，$(x_1,\cdots,x_n)$ が G_n' に落ちるとき（すなわち $z_1+\cdots+z_n\geqslant\log k_1$ のとき），$z_1+\cdots+z_n$ に $\log k_1$ を与えれば (15.4.49) および (15.4.50) の不等式は等式になり，(15.4.41) は $\mathscr{E}(n|\theta_0)$, $\mathscr{E}(n|\theta_1)$ に対して等式となる．このように上述の規則に応じて $z_1+\cdots+z_n$, $n=1,2,\cdots$ に $\log k_0$, $\log k_1$ を与えるならば，この修正された逐次確率比検定の場合には，等式 (15.4.42) が実際に実現されることがわかる．したがってこの条件の下では $\mathscr{H}_1$ に対して $\mathscr{H}_0$ を検定するとき，この修正された確率比検定よりも有効な逐次検定は存在しない．逐次確率比検定の最適な特性に関するこのような発見的な議論は Wald (1945) によってはじめて行なわれた．後に Wald と Wolfowitz (1948) は，より完全に，しかも厳密に論じている．

(g) 逐次確率比検定の停止

逐次確率比検定が $n=1,\cdots,N-1$ に対して停止しないならば，x_N の抜き取りに基づいて検定を停止させるために次の規則を採用するとしよう．

(15.4.52)
$$\log k_0<z_1+\cdots+z_N\leqslant 0$$

ならば $\mathscr{H}_0$ を採択する．

(15.4.53)
$$0<z_1+\cdots+z_N<\log k_1$$

ならば $\mathscr{H}_1$ を採択する．

この停止させられた逐次検定を S_N で表わし，α_N および β_N を S_N に関する第1種および第 2 種の誤りとし，α_N, β_N の上界を定める問題を考えよう．$G_{(N)}^\circ$, $G_{(N)}'$ を検定 S_N によってそれぞれ $\mathscr{H}_0$, $\mathscr{H}_1$ が採択される R_N の事象とすれば，$G_{(N)}^\circ$ と $G_{(N)}'$ は和が R_N の排反な事象になる．

便宜上 $G_{(N)}^\circ$, $G_{(N)}'$ を R_∞ の筒集合とみなせば，もちろん排反で，和が R_∞ で]

(15.4.54)
$$P(G_{(N)}'|\theta_0)=\alpha_N$$

である．前と同様に（停止させられた）逐次検定 S において $\mathscr{H}_0$, $\mathscr{H}_1$ が採択される R_∞ の事象をそれぞれ G°, G' とする．このとき次式が成り立つ．

(15.4.55)
$$P(G'|\theta_0)=\alpha.$$

停止させられた場合に，$\mathscr{H}_1$ が採択され，停止させられない場合に棄却（$\mathscr{H}_0$ が採択）される R_∞ の事象を G'^* とすれば

(15.4.56)
$$G_{(N)}'\subset(G'\cup G'^*).$$

さて (15.4.53) を満たす R_∞ の事象を J で表わせば

$$G'^*\subset J$$

よって

(15.4.57)
$$G_{(N)}'\subset(G'\cup J).$$

いま，$z_1+\cdots+z_N$ が漸近的に正規分布に従うように N を十分大きく選べば

(15.4.58)
$$P(0<z_1+\cdots+z_N<\log k_1|\theta_0)=\Phi(y_0')-\Phi(y_0)+O\left(\frac{1}{\sqrt{N}}\right)$$

が成り立つ．ただし

(15.4.59)
$$y_0=-\sqrt{N}\mathscr{E}(z|\theta_0)/\sigma(z|\theta_0)$$
$$y_0'=\sqrt{N}\left[\frac{1}{N}\log k_1-\mathscr{E}(z|\theta_0)\right]\Big/\sigma(z|\theta_0)$$

で，$\Phi(y)$ は $N(0,1)$ の c.d.f.．一方 $\mathscr{E}(z|\theta_0)$ と $\sigma^2(z|\theta_0)$ は確率変数

$$z=\log\left(\frac{f(x;\theta_1)}{f(x;\theta_0)}\right)$$

の p.d.f. $f(x;\theta_0)$ から定まる平均と分散である．

したがって $1/\sqrt{N}$ 次の項を無視すれば，$\alpha+\Phi(y_0')-\Phi(y_0)$ は α_N に対する上界である．同様に $1/\sqrt{N}$ 次の項を除外すれば，$\beta+\Phi(y_1')-\Phi(y_1)$ は β_N の上界になる．ただし

5.4.59 a)
$$y_1=\sqrt{N}\left[\frac{1}{N}\log k_0-\mathscr{E}(z|\theta_1)\right]\Big/\sigma(z|\theta_1)$$
$$y_1'=-\sqrt{N}\mathscr{E}(z|\theta_1)/\sigma(z|\theta_1)$$

で $\mathscr{E}(z|\theta_1)$ と $\sigma^2(z|\theta_1)$ は $f(x;\theta_1)$ より定まる z の平均と分散である．

α, β が小さいとき (y_1 における) 未知の k_0 と (y_0' における) 未知の k_1 は, それぞれ $\beta/(1-\alpha)$ と $(1-\beta)/\alpha$ で十分に近似される. これは 15.4(c) で示されている通りである.

要約すると次の Wald (1945) の結果が得られる.

15.4.9 確率変数

$$z = \log\left(\frac{f(x\,;\,\theta_1)}{f(x\,;\,\theta_0)}\right)$$

が $f(x\,;\,\theta_0)$, $f(x\,;\,\theta_1)$ のどちらから計算しても有限な平均と分散を持つならば, 停止させられた逐次確率比検定 S_N の第 1 種の誤り α_N の上界は

$$\alpha + \Phi(y_0') - \Phi(y_0) + O\left(\frac{1}{\sqrt{N}}\right)$$

である. ただし α は停止させられていない逐次検定 S に対する第 1 種の誤りで, y_0 と y_0' は (15.4.59) で与えられている. 同様に第 2 種の誤り β_N の上界は

$$\beta + \Phi(y_1') - \Phi(y_1) + O\left(\frac{1}{\sqrt{N}}\right)$$

である. ただし β は S の第 2 種の誤りである. 一方 y_1 と y_1' は (15.4.59a) で与えられている.

15.5 逐次確率比検定の 2 項分布への応用

15.4 節の結果を例証するために 2 項分布 $Bi(1,\theta)$ からの標本抽出の場合を考えよう. この場合 $(x_1, x_2, \cdots)$ は p.f.

(15.5.1) $$f(x\,;\,\theta) = \theta^x(1-\theta)^{1-x}, \qquad x = 0, 1$$

を持つ独立な確率変数列である. $\sum_{t=1}^{n} x_t$ を n_1 で表わせば

(15.5.2) $$\frac{Q_{1n}}{Q_{0n}} = \left(\frac{\theta_1}{\theta_0}\right)^{n_1}\left(\frac{1-\theta_1}{1-\theta_0}\right)^{n-n_1}$$

となる. k_0 と k_1 に対する近似値 (15.4.19) を用いて, a_n, b_n を次の式で定義する.

(15.5.3)
$$a_n = \frac{\log\left(\frac{\beta}{1-\alpha}\right) - n\log\left(\frac{1-\theta_1}{1-\theta_0}\right)}{\log\left[\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right]}$$
$$b_n = \frac{\log\left(\frac{1-\beta}{\alpha}\right) - n\log\left(\frac{1-\theta_1}{1-\theta_0}\right)}{\log\left[\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right]}.$$

このときの (15.4.1) の 3 つの不等式はそれぞれ

(15.5.4) $$n_1 \leqslant a_n, \qquad n_1 \geqslant b_n, \qquad a_n < n_1 < b_n$$

に帰着する. よって逐次過程は $a_n < n_1 < b_n$ であるかぎり続行し, $n_1 \leqslant a_n$ ならば x_n を抜き取り $\mathscr{H}_0$ を採択して停止し, $n_1 \geqslant b_n$ ならば x_n を抜き取り $\mathscr{H}_1$ を採択 ($\mathscr{H}_0$ を棄却) して停止する.

(15.4.37) において k_0, k_1 の近似値 (15.4.19) を用いると, 逐次検定の検査特性関数の近似式

(15.5.5) $$L(\theta) \cong \frac{1-\left(\frac{1-\beta}{\alpha}\right)^{h(\theta)}}{\left(\frac{\beta}{1-\alpha}\right)^{h(\theta)} - \left(\frac{1-\beta}{\alpha}\right)^{h(\theta)}}$$

が得られる. ただし $h(\theta)$ は (15.4.32) に

(15.5.6) $$\sum_{x=0}^{1}\left[\frac{\theta_1^x(1-\theta_1)^{1-x}}{\theta_0^x(1-\theta_0)^{1-x}}\right]^h \theta^x(1-\theta)^{1-x} = 1$$

を適用して定義された θ の関数である. (15.5.6) を簡単にすると, θ と h の関数関係

(15.5.7) $$\theta = \frac{1-\left(\frac{1-\theta_1}{1-\theta_0}\right)^h}{\left(\frac{\theta_1}{\theta_0}\right)^h - \left(\frac{1-\theta_1}{1-\theta_0}\right)^h}$$

を得る. $h(\theta_0) = 1$, $h(\theta_1) = -1$ に注意しよう. さらに (15.5.5) より, $L(\theta)$ の近似式は $L(\theta_0)$, $L(\theta_1)$ の正しい値を与えていることがわかる.

(15.4.40) に戻って, $\mathscr{E}(z|\theta)$ の値を計算すれば, 平均検査個数 $\mathscr{E}(n|\theta)$ は近似的に

(15.5.8) $$\mathscr{E}(n|\theta) \cong \frac{L(\theta)\log\left(\frac{\beta}{1-\alpha}\right) + (1-L(\theta))\log\left(\frac{1-\beta}{\alpha}\right)}{\theta\log\frac{\theta_1}{\theta_0} + (1-\theta)\log\left(\frac{1-\theta_1}{1-\theta_0}\right)}$$

で与えられる. よって (15.5.5) による $L(\theta)$ の (近似) 式および (15.5.7) の θ を (15.5.8) に代入すれば $\mathscr{E}(n|\theta)$ はパラメータ h の関数で表わされる. $\mathscr{E}(n|\theta)$ は $\theta = 0$,

θ_1, θ_2 のとき容易に求められる.

ワルドの逐次確率比検定を特定の分布からの抜き取りに適用した例が Wald (1947 *a*) および Statistical Research Group, Columbia University (1945, 1947) に多く発表されている.

15.6 逐次推定

(a) 概　　説

これまでパラメータの逐次推定に関する一般論は展開されていなかった. Wald(1947 *a*) は区間による逐次推定の一般的な問題に対する 1 つの定式化を与えたが, これを解いていない. しかし Stein と Wald (1947) は分散既知なる正規分布の平均に対して, 固定された長さと信頼係数を持つ信頼区間を求めるために, 逐次過程を考察している. 他の個々の推定問題も逐次推定と同様な目的で考察されている. この問題について多くの結果が Anscombe (1953) によって与えられている.

一般にパラメータの逐次推定の概念は次のように標本を抽出することである. すなわち推定される未知の母集団パラメータに依存しない意味で, あらかじめ定められた精度を持つ推定が得られるように, 標本を抽出することである. 精度の表わし方には種々の方法を設定することができる. そのうち簡単な方法として, 分散がパラメータに依存しないように推定量を与えることが考えられる. この方法はあらかじめ固定した大きさの標本を用いることによって行なわれることもある. 大標本の場合には, ある正則条件の下でこの問題の近似解が 12.3(e) 節で示されている.

精度の表現方法としては, 他にパラメータに対するもので, 信頼区間が標本抽出を行なう前に特定の長さを持つようにすることがある. この方法は大きさが固定された標本によるような特殊な場合にも用いられる. たとえば $N(\mu, \sigma^2)$ において μ を未知, σ^2 を既知とし, $\bar{x}$ をこの分布からの大きさ n の標本平均とする. このとき $n-1$ が $y_\gamma^2\sigma^2/\delta^2$ 以下の最大整数 ($y_\gamma^2\sigma^2/\delta^2$ が整数でないときは, この整数として n が選ばれる) として選ばれるならば, $\bar{x} \pm \frac{1}{2}\delta$ は μ の推定に対する (信頼) 係数 ($\geqslant \gamma$) を持つ長さ δ の信頼区間である. ただし y_γ は $\Phi(y_\gamma) - \Phi(-y_\gamma) = \gamma$ を満たし, $\Phi(y)$ は $N(0,1)$ の c.d.f. である.

(b) 正規分布の平均に対するスタインの固定された区間推定量

σ^2, μ がともに未知の場合, Stein (1945) は μ の推定に対して, 信頼係数 $\geqslant \gamma$ を持ち固定された長さ δ の信頼区間を示すために, 2 回抜き取り方式がどのように用いられるかを示した. 彼の結果は次のように述べられる.

15.6.1　$(x_1, \cdots, x_n)$ と $(x_{n+1}, \cdots, x_{n+m})$ はともに $N(\mu, \sigma^2)$ からの最初と次の独立な標本とする. s^2 を最初の標本の標本分散, $\bar{x}$ を合わせた標本の標本平均とする. $k = [2st_{n-1,\gamma}/\delta]^2$ とする. ただし $t_{n-1,\gamma}$ はスチューデント分布 $S(n-1)$ の右側 $100\left(1-\frac{1}{2}\gamma\right)\%$ 点である. m を $k-n \leqslant 0$ なら 0, $k-n>0$ なら $\geqslant k-n$ なる最小の正整数とすれば, $\bar{x} \pm \frac{1}{2}\delta$ は信頼係数 $\geqslant \gamma$ を持つ μ に対する長さ δ の信頼区間である.

15.6.1 を示すには次の事柄がわかれば容易である. すなわち m が確率変数のときでも, $N(\mu, \sigma^2)$ からの大きさ m の標本平均 $\bar{x}$ に対して, $(\bar{x}-\mu)/[\sigma/\sqrt{m+n}]$ が分布 $N(0,1)$ を持ち, カイ 2 乗分布 $C(n-1)$ に従う $(n-1)s^2/\sigma^2$ と独立である. これが示されれば $(\bar{x}-\mu)\sqrt{m+n}/s$ はスチューデント分布 $S(n-1)$ を持つ. ゆえに

$$P\left(-t_{n-1,\gamma} < \frac{(\bar{x}-\mu)}{s}\sqrt{m+n} < +t_{n-1,\gamma}\right) = \gamma. \tag{15.6.1}$$

これは

$$P\left(\bar{x} - \frac{st_{n-1,\gamma}}{\sqrt{m+n}} < \mu < \bar{x} + \frac{st_{n-1,\gamma}}{\sqrt{m+n}}\right) = \gamma \tag{15.6.2}$$

と同値である.

$st_{n-1,\gamma}/\sqrt{n} \leqslant \frac{1}{2}\delta$ ならば, $\bar{x} \pm \frac{1}{2}\delta$ は最初の標本だけの平均で与えられた μ に対する長さ δ の信頼区間である. この μ は信頼係数 $\geqslant \gamma$ を持つ. しかし $st_{n-1,\gamma}/\sqrt{m} > \frac{1}{2}\delta$ ならば, 次に大きさ m の標本を抜き取る. ただし m は

$$\frac{st_{n-1,\gamma}}{\sqrt{m+n}} \leqslant \frac{1}{2}\delta.$$

すなわち

$$m \geqslant \left[\frac{2st_{n-1,\gamma}}{\delta}\right]^2 - n$$

なる最小の正整数である. この場合も $\bar{x} \pm \frac{1}{2}\delta$ は信頼区間 $\geqslant \gamma$ を持つ 2 標本を合わせた標本の平均 μ に対する長さ $\leqslant \delta$ の信頼区間になる. これで **15.6.1** が示された.

問　　題

15.1　$(x_1, x_2, \cdots)$ が有限区間 (a, b) で定義された等しい平均 $\mathscr{E}(x)$ を持つ独立な確率変数で，$\mathscr{E}(n)$ が有限ならば **15.2.1** が成り立つことを示せ．

15.2　$(x_1, x_2, \cdots)$ は c.d.f. $F_1(x\,;\theta), F_2(x\,;\theta), \cdots$ を持つ独立な確率変数列とする．$E^{(i)}$ を

$$\int_{E^{(i)}} dF_i(x\,;\theta) = p_i(\theta) \leqslant p < 1$$

となる集合とする．S が次で定義された実験続行集合列 $G_1, G_2, \cdots$ を持つ任意の逐次過程ならば，確率 1 で停止することを示せ．

$$G_n = E^{(1)} \times E^{(2)} \times \cdots \times E^{(n)}.$$

15.3　独立な確率変数列 $(x_1, x_2, \cdots)$ は p.d.f. $\theta e^{-\theta x}$, $x > 0$, $\theta > 0$ を持つ母集団から抜き取られたとする．仮説 $\mathscr{H}_1 : \theta = \theta_1$ に対して仮説 $\mathscr{H}_0 : \theta = \theta_0$（ただし $\theta_1 > \theta_0$）の検定に対するカルテシアン逐次過程を構成したい．ただし第 1 種の誤り，第 2 種の誤りをそれぞれ α, β とする．確率比基準により与えられる G_0°, G_0', G_0 は区間 (x_2, ∞), $(0, x_1)$, (x_1, x_2) になることを示せ．ただし $x_1 < x_2$ は条件式

$$\frac{e^{-\theta_0 x_2}}{1 - e^{-\theta_0 x_1}} = \frac{1-\alpha}{\alpha}, \qquad \frac{e^{-\theta_1 x_2}}{1 - e^{-\theta_1 x_1}} = \frac{\beta}{1-\beta}$$

を満たす．この逐次過程に対する平均検査個数 $\mathscr{E}(n|\theta)$ は $\theta = \dfrac{\log x_2 - \log x_1}{x_2 - x_1}$ のとき最大になることを示せ．

15.4（続き）$\mathscr{H}_1$ に対する $\mathscr{H}_0$ の仮説検定に対して，r 重カルテシアン逐次検定を用いるとする．確率比基準で定められる（r 次元ユークリッド空間における）G_0°, G_0', G_0 は次のように定義されることを示せ．

$$G_0^\circ : \left\{(x_1, \cdots, x_r) : \sum_1^r x_i \geqslant y_2\right\}$$

$$G_0' : \left\{(x_1, \cdots, x_r) : \sum_1^r x_i \leqslant y_1\right\}$$

$$G_0 : \left\{(x_1, \cdots, x_r) : y_1 < \sum_1^r x_i < y_2\right\}.$$

ただし (y_1, y_2) は

$$\int_{y_2}^\infty f(y\,;\theta_0)\,dy = \frac{1-\alpha}{\alpha}\int_0^{y_1} f(y\,;\theta_0)\,dy$$

$$\int_{y_2}^\infty f(y\,;\theta_1)\,dy = \frac{\beta}{1-\beta}\int_0^{y_1} f(y\,;\theta_1)\,dy$$

および

$$f(y\,;\theta) = \frac{\theta^r y^{r-1} e^{-\theta y}}{\Gamma(r)}, \qquad y > 0$$

を満たす．

15.5　15.3(a) 節で議論されたカルテシアン逐次検定において，過程が n 回目の試行で停止しないならば，この過程は $\mathscr{H}_0$ か $\mathscr{H}_1$ を選んで，それぞれ確率 $a/(a+b)$, $b/(a+b)$ で停止するとする（勝手に打ち切られる）．このとき

$$L_n(\theta) = \frac{a}{a+b}, \qquad M_n(\theta) = \frac{b}{a+b}, \qquad \mathscr{E}(n|\theta) = \frac{1-c^n}{a+b}$$

および第 1 種の誤り，第 2 種の誤り α, β は (15.3.8) を満たすことを示せ．

15.6　G_0° を (15.3.10)，G_0' を (15.3.11) で定義したカルテシアン逐次検定は $a_0/(a_0+b_0)$ を α, $\mathscr{E}(n|\theta_0)$（すなわち $1/(a_0+b_0)$）を n_0 に固定した他の任意のカルテシアン検定よりも強いことを証明せよ．

15.7　$(x_1, x_2, \cdots)$ はすべて正規分布 $N(\theta, 1)$ を持つ独立確率変数列とする．$\mathscr{H}_0$ は仮説 $\theta = \theta_0$，$\mathscr{H}_1$ は仮説 $\theta = \theta_1$（ただし $\theta_1 > \theta_0$）とする．第 1 種の誤り，第 2 種の誤りを α, β とし，k_0 と k_1 の近似値として (15.4.19) で与えられる値をとる．$\mathscr{H}_1$ に対する $\mathscr{H}_0$ の検定の逐次確率比検定においては，はじめて

$$x_1 + \cdots + x_n \leqslant \frac{1}{(\theta_1 - \theta_0)} \log\left(\frac{\beta}{1-\alpha}\right) + n\frac{\theta_1 + \theta_2}{2}$$

となる n 回目の抜き取りに基づいて $\mathscr{H}_0$ が採択され，はじめて

$$x_1 + \cdots + x_n \geqslant \frac{1}{(\theta_1 - \theta_0)} \log\frac{1-\beta}{\alpha} + n\frac{\theta_1 + \theta_2}{2}$$

となる n 回目の抜き取りに基づいて $\mathscr{H}_1$ が採択され，$x_1 + \cdots + x_n$ が上記の 2 つの間に落ちるならば，$(n+1)$ 回目の抜き取りを行なうことを示せ．

また

$$L(\theta) \cong \frac{\left(\dfrac{1-\beta}{\alpha}\right)^h - 1}{\left(\dfrac{1-\beta}{\alpha}\right)^h - \left(\dfrac{\beta}{1-\alpha}\right)^h} \qquad \left(\text{ただし } h = \frac{\theta_1 + \theta_0 - 2\theta}{\theta_1 - \theta_0}\right)$$

で，$L(\theta)$ に対するこの近似式は第 1 種の誤り，第 2 種の誤りで与えられたように点 $(\theta_0, 1-\alpha)$ と (θ_1, β) を通ることを示せ．さらに平均検査個数は

$$\mathscr{E}(n|\theta) \cong -2\frac{L(\theta)\log\left(\dfrac{\beta}{1-\alpha}\right) + (1 - L(\theta))\log\left(\dfrac{1-\beta}{\alpha}\right)}{h(\theta_1 - \theta_0)^2}$$

であることを示せ．これらの結果を σ^2 が既知の $N(\theta, \sigma^2)$ からの標本抽出の場合に拡張せよ．

15.8　$(x_1, x_2, \cdots)$ は正規分布 $N(0, \sigma^2)$ を持つ確率変数列とする．$\mathscr{H}_0$ を仮説 $\sigma^2 = \sigma_0^2$，$\mathscr{H}_1$ を仮説 $\sigma^2 = \sigma_1^2$（ただし $\sigma_1^2 > \sigma_0^2$）とする．第 1 種の誤り α，第 2 種の誤り β に対して (15.4.19) での k_0, k_1 の近似値を用いると，$\mathscr{H}_1$ に対する $\mathscr{H}_0$ の検定の逐次確率

比検定は次のようになることを示せ．はじめて

$$x_1^2+\cdots+x_n \leqslant \frac{\sigma_0^2\sigma_1^2\left[2\log\left(\dfrac{\beta}{1-\alpha}\right)+n\log\left(\dfrac{\sigma_1^2}{\sigma_0^2}\right)\right]}{\sigma_1^2-\sigma_0^2}$$

となる n 回目の試行に基づいて $\mathscr{H}_0$ を採択し，はじめて

$$x_1^2+\cdots+x_n^2 \geqslant \frac{\sigma_0^2\sigma_1^2\left[2\log\left(\dfrac{1-\beta}{\alpha}\right)+n\log\left(\dfrac{\sigma_1^2}{\sigma_0^2}\right)\right]}{\sigma_1^2-\sigma_0^2}$$

となる n 回目の試行に基づいて $\mathscr{H}_1$ を採択し，そして $x_1^2+\cdots+x_n^2$ が上記の 2 つの値の間に落るならば $(n+1)$ 回目の観察を行なう．

15.9 (続き)　検査特性関数 $L(\sigma^2)$ は近似的に

$$L(\sigma^2)\cong\frac{\left(\dfrac{1-\beta}{\alpha}\right)^h-1}{\left(\dfrac{1-\beta}{\alpha}\right)^h-\left(\dfrac{\beta}{1-\alpha}\right)^h}$$

で与えられることを示せ．ただし σ^2 と h には次の関数関係がある．

$$\sigma^2=\frac{\sigma_0^2\sigma_1^2\left[1-\left(\dfrac{\sigma_0}{\sigma_1}\right)^{2h}\right]}{h(\sigma_1^2-\sigma_0^2)}.$$

また $L(\sigma^2)$ に対するこの近似式のグラフは 2 点 $(\sigma_0^2, 1-\alpha)$，(σ_1^2, β) を通ることを確かめよ．

さらに平均検査個数は

$$\mathscr{E}(n\,|\,\sigma^2)\cong\frac{2\sigma_0^2\sigma_1^2\left[L(\sigma^2)\log\left(\dfrac{\beta}{1-\alpha}\right)+(1-L(\sigma^2))\log\left(\dfrac{1-\beta}{\alpha}\right)\right]}{(\sigma_1^2-\sigma_0^2)\sigma^2+\sigma_0^2\sigma_1^2\log(\sigma_0^2/\sigma_1^2)}$$

で与えられることを示せ．

15.10　対立仮説 $\mathscr{H}_1$《大きさ n の標本はポアソン分布 $Po(\theta_1)$ から抽出される》に対する，仮説《この標本はポアソン分布 $Po(\theta_0)$ から抽出される》(ただし $\theta_1>\theta_0$) の逐次確率比検定を定めよ．この検定の検査特性関数 $L(\theta)$ および平均検査個数の近似式を見い出せ．

15.11　2 項待ち時間分布

$$\binom{n-1}{k-1}p^k(1-p)^{n-k},\qquad n=k,k+1,\cdots$$

において $\tilde{p}=\dfrac{k-1}{n-1}$ とする．$\tilde{p}$ は p に対する不偏推定量で

$$\sigma^2(\tilde{p})=\frac{p^2}{k-2}\left[1-\left(\frac{k-1}{k-3}\right)p+\frac{2(k-1)p^2}{(k-3)(k-4)}-\cdots\right]$$

したがって $\tilde{p}$ の変動係数，すなわち $\sigma(\tilde{p})/\mathscr{E}(\tilde{p})$ は

$$\frac{1}{\sqrt{k-2}}(1+O(p)).$$

であることを示せ．[Haldane (1945) を参照せよ]．

15.12　s^2 は $N(\mu,\sigma^2)$ からの大きさ n の標本の標本分散とする．任意の $\varepsilon>0$ に対して，m を $\geqslant(n-1)s^2/[(n-3)\varepsilon]$ となる最小の整数とする．$\bar{x}'$ を $N(\mu,\sigma^2)$ からの大きさ m の独立な標本の標本平均とすれば，$\sigma^2(\bar{x}')\leqslant\varepsilon$ であることを示せ．

15.13　s_1^2，s_2^2 はそれぞれ $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$ からの大きさ n_1, n_2 の独立な標本の標本分散とする．任意の ε (>0) に対して，m_1 を $\geqslant[2(n_1-1)s_1^2]/[(n_1-3)\varepsilon]$ となる最小の整数とし，また m_2 を $\geqslant[2(n_2-1)s_2^2]/[(n_2-3)\varepsilon]$ となる最小の整数とする．さらに $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$ のおのおのからの大きさ m_1, m_2 の標本は独立で，$\bar{x}_1$, $\bar{x}_2$ をおのおのの標本平均とする．このとき $\sigma^2(\bar{x}_1-\bar{x}_2)\leqslant\varepsilon$ を示せ．

第16章　統計的決定関数

16.1 概　　説

ネイマンやピヤソンによって展開された検定論において，仮説を誤って受け入れたり，拒否したりするときの危険性が認識され，強調されるようになった．後に1冊の本として発表された最初の2つの論文の中で，Wald (1947 *b*, 1949 *a*, 1950) はこのような危険性に関する理論をより広汎な統計的な問題へと拡張し，いわゆる統計的決定関数の理論を展開した．その展開は，Neumann と Morgenstern (1944) によって定式化されたゲームの理論とか，統計的処理の際に起る危険性の評価に関する一般論を構築したいという希望とかが，いくぶんか刺激になっている．統計的決定関数に関するワルドのオリジナルな仕事は多くの論文やいくつかの著作に見られる．たとえば，離散型標本空間に対する統計的決定理論を取り扱っている Blackwell と Girshick (1954) は比較的理解し易いものである．この章では，最も簡単な場合について統計的決定関数論での主要な概念および基本的な結果をいくつか簡単に紹介しよう．一般論まで深入りしないことにする．一般論および，さらに詳細な結果に興味ある読者は，ワルド，ブラックウェルとガーシックの著作を参照されたい．さらに，初等的立場で書かれてある本としては，Chernoff と Moses (1959), Luce と Raiffa (1957), Raiffa と Schlaifer (1961), Savage (1954), Weiss (1961) などがある．

16.2 定 義 と 用 語

x は標本空間 R を持つ離散型確率変数，θ は有限個の点 $\theta_1, \cdots, \theta_h$ からなるパラメータ空間 Ω の点とする．いま，$p(x|\theta)$ を θ が与えられたときの x の確率分布とすると，h 個の確率分布 $p(x|\theta_1), \cdots, p(x|\theta_h)$ が考えられる．x に関して観測すると，つねにこのうちの1つが“真の”分布であると仮定されている．

a は x の観測結果によって，θ に関してなされ得る決定（あるいは結論）の集合 A の任意の1点であるとする．たとえば，$\theta_1, \cdots, \theta_h$ を実数で，c をある臨界値とすれば，A としては次の2点 a_1, a_2 からなる場合が考えられる．ここに，a_1 は $\theta \leqq c$ なる決定，a_2 は $\theta > c$ なる決定である．もちろん，A は任意有限個の点からなってもよいし，可算無限個の点集合でもよい．A は**決定空間**と呼ばれている（**結論空間**と呼ぶのは，性急な用語という感じがする）．x の観測がどの決定 $a \in A$ がなされるかということに何らかの関係があるならば，a は x に依存しなければならない．よって，標本点 x が起きたときに A のどの決定 a が下されるかを定める**決定関数** $d(x)$ が考えられる．このように，$d(x)$ は標本空間 R の各点 x に対して，ある $a \in A$ を値に持つ x の1価関数である．考えられる任意の決定関数 $d(x)$ は，ある決定関数のクラス D の要素である．決定関数 $d(x)$ および有限個の点からなる標本空間 R が与えられたときには，A の要素も有限になることを注意しておこう．

さて，各 $\theta \in \Omega$ に対して任意の決定 $a \in A$ をとることが可能である．θ が真のとき，決定 a の結果が $L(\theta, a)$ という損失あるいは費用を招くとすれば，有界1価実数値関数 $L(\theta, a)$ が直積空間 $\Omega \times A$ の各点 (θ, a) で定義される．これは**損失関数**と呼ばれている．任意の決定関数 $d \in D$ とパラメータ $\theta \in \Omega$ に対して，**危険関数** $r(\theta, d)$ が損失関数 $L(\theta, d(x))$ の全標本空間にわたる平均値として定義される．すなわち，

$$r(\theta, d) = \mathscr{E}[L(\theta, d(x))] = \sum_{x \in R} L(\theta, d(x)) p(x|\theta). \tag{16.2.1}$$

このように，$r(\theta, d)$ は直積空間 $\Omega \times D$ の各点で定義された有界な関数である．なぜなら，$L(\theta, d(x))$ は有界であるからである．

注意　これまでに述べたいくつかの概念をゲームの理論の初歩的な概念を用いて説明しておくと便利である．自然をプレーヤーI，統計家をプレーヤーIIとする．自然の戦略あるいは選択は $\theta_1, \cdots, \theta_h$，一方，統計家のそれは D に含まれる d である．このようにすると，自然によって任意の θ が，統計家によって任意の d が選ばれ，x が起ったとき，$L(\theta, d(x))$ が統計家の損失（自然の支払い額）になる．(16.2.1) で定義された量 $r(\theta, d)$ は，戦略として $d(x)$ を採用した統計家の平均損失である．

いま，ある損失関数 $L(\theta, a)$ と決定関数 d が与えられ，これに危険関数 $r(\theta, d)$ が対応しているとしよう．D のある決定関数よりも好ましい別の決定関数を選ぶ場合に，どのような基準をもってすればよいだろうかという問題が考えられる．好都合なことに危険関数そのものが1つの正当な評価基準を与えている．d と d^* を D の2つの決定関数とする．危険関数に基づいて d と d^* を比較して，すべての $\theta \in \Omega$ に対して

(16.2.2)
$$r(\theta, d^*) \leqslant r(\theta, d)$$

かつ，少なくとも1つの $\theta \in \Omega$ に対して

(16.2.3)
$$r(\theta, d^*) < r(\theta, d)$$

が成り立つとき，d^* は d よりも好ましいということになるだろう．(16.2.2) かつ (16.2.3) が成り立つとき，d^* は d よりも一様によい決定関数と呼ばれる．D のどの決定関数も d^* より一様によい決定関数でないとき，d^* を許容的決定関数と呼ぶ．決定関数のクラス D に含まれない任意の d に対して，d より一様によい d^* が D の中に存在するとき，D は完備類といわれる．

決定関数の完備類 D が完備な（真の）部分クラスを含まないとき，D を最小完備類という．統計的決定関数の一般論では決定関数の完備類の概念は重要な役割を果たしている．決定関数の完備類と最小完備類との関係は次の定理 **16.2.1**, **16.2.2** で与えられている．

16.2.1　最小完備類が存在するならば，許容的決定関数のクラスに等しい．

D_0 を許容的決定関数のクラスとする．いま，D を最小完備類とすれば，D_0 は D の部分集合である．さて，d' を D_0 にはいっていない D の要素とする．したがって，d' より一様によい d'' が存在する．しかし，D は最小完備類だから，d'' は D の要素ではない．よって，d'' より一様によい要素 $d''' \in D$ が存在する．したがって，d''' は d' より一様によい．これは D が最小完備類であるから不可能である．すなわち，D_0 は D の真部分集合でない．ゆえに，D と D_0 は一致する．

許容的決定関数のすべてからなるクラス D_0 が完備ならば，D_0 は最小完備類である．この事実と **16.2.1** より次の事柄が成り立つ．

16.2.2　決定関数の最小完備類が存在するための必要十分条件は，許容的決定関数のクラスが完備なことである．

Wald (1950) は，一般的な条件の下で許容的決定関数の集合が完備類を構成すること，特に，ベイズ解（後に考察する）全体の集合は完備であることなどを示している．

16.3　決定問題のミニマックス解

危険関数 $r(\theta, d)$ に戻ろう．任意に与えられた決定関数 d および損失関数 $L(\theta, a)$ に対して，θ の可能な選択，すなわち $\theta_1, \cdots, \theta_h$ に対応する危険ベクトルは $(r(\theta_1, d), \cdots, r(\theta_h, d))$ である．さて，このベクトルの最大要素を考える．この要素は d の関数である．最大要素を最小にするクラス D の決定関数を d^* とする．いい換えれば，d^* は D のすべての d に対して

(16.3.1)
$$\underset{\theta \in \Omega}{\text{l.u.b.}}\ r(\theta, d^*) \leqslant \underset{\theta \in \Omega}{\text{l.u.b.}}\ r(\theta, d)$$

となる D の要素である．このとき，d^* または d^* に対応する危険ベクトル $(r(\theta_1, d^*), \cdots, r(\theta_h, d^*))$ を決定問題のミニマックス解と呼ぶ．

D が有限個の要素 d からなるとき，少なくとも1つのミニマックス解 d^* が存在し，対応する危険（値）が

(16.3.2)
$$\min_{d \in D} \{\max_{\theta \in \Omega} r(\theta, d)\}$$

であることは明らかである．ミニマックス危険（値）を与える d（あるいは複数個の d）は (16.3.2) で示された順序に従って $\Omega \times D$ の有限集合上で $r(\theta, d)$ の値をすべて挙げることによって見い出される．

しかし，D が無限個の要素を含むときは，立場が相当複雑になる．この場合，幾何学的表現が役立つだろう．各 $d \in D$ に対して，h 個の成分の組 $(r(\theta_1, d), \cdots, r(\theta_h, d))$ はユークリッド空間 R_h の点 r として表わせる．すべての $d \in D$ に対応するこのような点 r の集合は R_h の部分集合 E を構成する．危険関数 r は $\Omega \times D$ 上で有界だから E は有界である．こうすれば，D に無限個の要素があるときは，決定問題のミニマックス解を求める問題は，E の各点 r の最大要素を見い出して，この最大要素を最小にすることと同値である．解 d^* が存在すれば，それは次の方法で表わすことができる．E_i を $r(\theta_i, d) \geqslant \max_{j \neq i} \{r(\theta_j, d)\}$ となる E の部分集合とする．ただし，$i = 1, \cdots, h$．このとき，$E = E_1 \cup E_2 \cup \cdots \cup E_h$ が成立する．しかし $E_1, E_2, \cdots, E_h$ は必ずしも排反でない．

なぜなら，2 つないし 3 つ以上の等しい最大の座標を持つ E の任意の点は，それぞれ $E_1, \cdots, E_h$ の 2 つないし 3 つ以上に属すかもしれないからである．さて，各 i $(i=1, \cdots, h)$ について，M_i を E_i のすべての点の i 番目の座標の最大下界とする．このとき D の点 d^* (E の点) が

(16.3.3)
$$\max_i r(\theta_i, d^*) = \min\{M_1, \cdots, M_h\}$$

を満たすならば，d^* はミニマックス解である．E が閉集合，すなわち，E の任意の点列の極限点が E に属すならば，少なくともこのような d^* が 1 つ存在する．

ある弱い条件の下で，E が閉集合であることを示そう．まず，$D \times R$ の任意の点に対して，$(L(\theta_1, d(x)), \cdots, L(\theta_h, d(x)))$ はユークリッド空間 R_h の点 L になることに注意しよう．F をこのような L の点集合とすれば，$L(\theta, d(x))$ が有界だから，F は有界集合である．次の **16.3.1** を示そう．

16.3.1　　F が閉集合ならば，E も閉集合である．

最初に，R が有限個の点を含んでいる場合を考える．このとき，A は有限個の点を持つ．$d(x)$ は 1 価だから，$L(\theta_i, d(x))$ は有限個の値をとる．したがって，標本空間 R にわたって $L(\theta_i, d(x))$ の値を $p(x|\theta_i)$, $i=1, \cdots, h$ で加重平均した $r(\theta_i, d)$ も有限個の値を持つ．ゆえに，R が有限個の点を含めば，E も有限個である．よって，E は閉集合である．したがって，この場合，決定問題のミニマックス解が存在する．

さて，R が（可付番）無限に多くの点を含む場合，この点を x_t, $t=1, 2, \cdots$ という具合に番号をつけておく．D の任意の d（したがって，E の対応する点 r）に対して，

(16.3.4)
$$r(\theta_i, d) = \sum_{t=1}^{\infty} L(\theta_i, d(x_t))p(x_t|\theta_i)$$

である．このように，d はユークリッド空間 R_h の点列

(16.3.5)
$$(L(\theta_1, d(x_t)), \cdots, L(\theta_h, d(x_t))), \qquad t=1, 2, \cdots$$

によって特徴づけられる．D の各 d に対するこのような点列全体が R_h の点集合 F を生成する．この F は有界である．なぜなら，$L(\theta, d(x))$ が有界であるからである．F を閉だと仮定しておいて，E の任意の収束点列 r_α, $\alpha=1, 2, \cdots$ を考えよう．ただし，r_α は座標

(16.3.6)
$$(r(\theta_1, d_\alpha), \cdots, r(\theta_h, d_\alpha))$$

を持つ．この点列の極限点を r^* とすれば，r^* は

(16.3.7)
$$(r(\theta_1, d^*), \cdots, r(\theta_h, d^*))$$

で表わされる．r^* もまた E に属すことを示せば，E が閉集合であることがわかる．さて，d_α に対応する F の点列は

(16.3.8)
$$(L(\theta_1, d_\alpha(x_t)), \cdots, L(\theta_h, d_\alpha(x_t))), \qquad t=1, 2, \cdots$$

である．$\alpha=1, 2, \cdots$ に対して，(16.3.8) は F における 2 重点列 $L_{\alpha t}$ となる．この 2 重点列に対してカントールの対角論法を用いると，列 d_α, $\alpha=1, 2, \cdots$ から部分列 $d_{\alpha\beta}$, $\beta=1, 2, \cdots$ が選ばれて，次のようになる．すなわち，F の点列

(16.3.9)
$$(L(\theta_1, d_{\alpha\beta}(x_t)), \cdots, L(\theta_h, d_{\alpha\beta}(x_t)))$$

は，各 t に対して，$\beta \longrightarrow \infty$ のとき

(16.3.10)
$$(L(\theta_1, d^*(x_t)), \cdots, L(\theta_h, d^*(x_t)))$$

に収束する．F は閉だから，(16.3.10) の点は，$t=1, 2, \cdots$ に対して，F の点である．さて，危険関数

(16.3.11)
$$r(\theta, d_{\alpha\beta}) = \sum_{t=1}^{\infty} L(\theta, d_{\alpha\beta}(x_t))p(x_t|\theta)$$

において，$\beta \longrightarrow \infty$ として極限をとる．(16.3.9) の点列は各 t について，(16.3.10) に収束するから

(16.3.12)
$$\lim_{\beta\to\infty} r(\theta, d_{\alpha\beta}) = \lim_{\beta\to\infty} \sum_{t=1}^{\infty} L(\theta, d_{\alpha\beta}(x_t))p(x_t|\theta)$$
$$= \sum_{t=1}^{\infty} \lim_{\beta\to\infty} L(\theta, d_{\alpha\beta}(x_t))p(x_t|\theta) = \sum_{t=1}^{\infty} L(\theta, d^*(x_t))p(x_t|\theta)$$

が成り立つ．lim と $\sum$ の交換は，損失関数が有界で，$p(x_t|\theta)$ が確率分布であることより可能である．(16.3.10) の点は F に属すから，$\sum_{t=1}^{\infty} L(\theta, d^*(x_t))p(x_t|\theta)$ は E の点である．すなわち，$r(\theta, d^*)$ は E に属す．ゆえに E は閉集合である．

このように，すべての $d \in D$ および $x \in R$ に対する R_h の点 $(L(\theta_1, d(x)), \cdots, L(\theta_h, d(x)))$ からなる集合 F が閉ならば，すべての $d \in D$ に対する R_h の点 $(r(\theta_1, d), \cdots, r(\theta_h, d))$ の集合 E も閉であり，統計的決定問題のミニマックス解 d^* が D において存在する．

このミニマックス手法を個々の問題に適用する場合，一般に，かなりの計算を要する．ここでは，非常に簡単な例を選んで，この方法を示そう．

例題　x は標本空間 $(0, 1)$ を持つ離散型確率変数で，分布

$$p(x|\theta_i) = \binom{1}{x}\theta_i^x(1-\theta_i)^{1-x}, \qquad \theta_1 = \frac{1}{4}, \theta_2 = \frac{1}{2}$$

を持つとする．決定空間 A は 2 点 a_1, a_2 からなるとする．損失関数 $L(\theta_i, a_j)$ は次のように定義されている．

	a_1	a_2
θ_1	1	4
θ_2	3	2

さて，各標本点 x に対して可能な決定は a_1 と a_2 である．各標本点で決定 a_1, a_2 のどちらかが許されるならば，標本空間上の次の 4 つの可能な決定関数 d_α, $\alpha = 1, 2, 3, 4$ が存在する．

$$\begin{aligned} d_1(0) &= a_1, & d_1(1) &= a_1 \\ d_2(0) &= a_1, & d_2(1) &= a_2 \\ d_3(0) &= a_2, & d_3(1) &= a_1 \\ d_4(0) &= a_2, & d_4(1) &= a_2 \end{aligned}$$

(16.2.1)，すなわち

$$r(\theta_i, d_\alpha) = \sum_{x=0}^{1} L(\theta_i, d_\alpha(x))p(x|\theta_i), \qquad i = 1, 2, \qquad \alpha = 1, \cdots, 4$$

を用いると，4 つの危険ベクトル $(r(\theta_1, d_\alpha), r(\theta_2, d_\alpha))$, $\alpha = 1, \cdots, 4$ が計算される．実際，$d_\alpha(x)$, $L(\theta_i, d_\alpha)$, $p(x|\theta_i)$ に数値を代入すると 4 つの危険ベクトルは $(1, 3)$, $\left(\frac{7}{4}, \frac{5}{2}\right)$, $\left(\frac{13}{4}, \frac{5}{2}\right)$, $(4, 2)$ である．4 つのベクトルの最大座標はそれぞれ 3, $\frac{5}{2}$, $\frac{13}{4}$, 4. この中の最小値は $\frac{5}{2}$ である．ゆえに，危険ベクトルが $(r(\theta_1, d_2), r(\theta_2, d_2))$ のとき，解となり，ミニマックス解は決定関数 $d_2(x)$ で，対応する危険ベクトルは $\left(\frac{7}{4}, \frac{5}{2}\right)$ である．

16.4　統計的決定問題のベイズ解

(a)　特定の事前分布に対する解

ここでは，θ の値はある事前確率分布 $q(\theta)$, $\theta = \theta_1, \cdots, \theta_h$ に応じて起ると仮定する．そうすると，D の与えられた各決定関数 d に対して，危険関数 $r(\theta, d)$ は θ に関する事前分布 $q(\theta)$ で平均されて，新しく平均危険 $\bar{r}(q, d)$ が考えられる．すなわち，$\bar{r}(q, d)$ は次のように定義される．

(16.4.1)
$$\bar{r}(q, d) = \sum_{i=1}^{h} r(\theta_i, d)q(\theta_i) = \sum_{t=1}^{\infty}\sum_{i=1}^{h} L(\theta_i, d(x_t))p(x_t|\theta_i)q(\theta_i).$$

一度，$q(\theta)$ を知った上で，D のすべての d に対して

(16.4.2)
$$\bar{r}(q, d^*) \leqq \bar{r}(q, d)$$

となる D の決定関数 d^* を見い出すことができたならば，d^* を統計的決定問題の最適解とみなしてもよかろう．このように $\bar{r}(q, d)$ を最小にする決定関数は特定の事前分布 $q(\theta)$ に関するベイズ解と呼ばれる．

$q(\theta_i) > 0$, $i = 1, \cdots, h$ で $q(\theta_1) + \cdots + q(\theta_h) = 1$ だから，$(q(\theta_1), \cdots, q(\theta_h))$ はユークリッド空間 R_h の h 個の点 $(1, 0, \cdots, 0), (0, 1, 0, \cdots, 0), \cdots, (0, \cdots, 0, 1)$ で張られる $(h-1)$ 次元単体 G 上の点として表わされる．

さて，決定者にとってある未知の**非退化事前分布**が与えられた場合を想定しよう．どのようにして，危険を最小にしたらよいだろうか？

$\bar{r}(q, d)$ は次で表わされる．すなわち

(16.4.3)
$$\bar{r}(q, d) = \sum_{t=1}^{\infty}\sum_{i=1}^{h} L(\theta_i, d(x_t))Q(\theta_i|x_t)P_q(x_t)$$

ただし

(16.4.4)
$$Q(\theta_i|x_t) = \frac{p(x_t|\theta_i)q(\theta_i)}{P_q(x_t)},$$

ここに

(16.4.5)
$$P_q(x_t) = \sum_{i=1}^{h} p(x_t|\theta_i)q(\theta_i).$$

$Q(\theta_i|x_t)$ は $x = x_t$ および事前分布 $q(\theta_i)$ が与えられたとき，$\theta = \theta_i$ となる事後確率である．

さて，$q(\theta)$ と t が与えられたとき，各 t に対して

(16.4.6)
$$\sum_{i=1}^{h} L(\theta_i, d^*(x_t))Q(\theta_i|x_t)$$

を

(16.4.7)
$$\sum_{i=1}^{h} L(\theta_i, d(x_t))Q(\theta_i|x_t)$$

の F のすべての点 $(L(\theta_1, d(x_t)), \cdots, L(\theta_h, d(x_t)))$ にわたる最大下界とする．もちろん，

t を固定すれば，$(Q(\theta_1|x_t), \cdots, Q(\theta_h|x_t))$ は G の点になっている．したがって，標本空間 R の点列 $x_1, x_2, \cdots$ は F におけるベクトル列および G における対応する点列を定めることになる．もし，F が各 t と D のすべての d に対し，ベクトル

(16.4.8) $$(L(\theta_1, d(x_t))Q(\theta_1|x_t), \cdots, L(\theta_h, d(x_t))Q(\theta_h|x_t))$$

および各 t について最小値をとるベクトル

(16.4.9) $$(L(\theta_1, d^*(x_t))Q(\theta_1|x_t), \cdots, L(\theta_h, d^*(x_t))Q(\theta_h|x_t))$$

を含むように拡大されるならば，E は閉じている．したがって D も閉じていて，(16.4.6) は (16.4.7) の表現に対して最小値を与えることになる．いい換えれば，D はこの中のすべての d に対して

(16.4.10) $$\sum_{t=1}^{\infty}\sum_{i=1}^{h} L(\theta_i, d^*(x_t))Q(\theta_i|x_t)P_q(x_t) \leqq \bar{r}(q, d)$$

となる d^* を少なくとも 1 つ含んでいる．このような d^* はどれも事前分布 $q(\theta)$ に対する決定問題のベイズ解になっている．このような d^* の集合は空でなく，閉じている．したがって，許容的決定関数 d^* のクラスは完備である．まとめると次のようになる．

16.4.1　　与えられた事前分布 $q(\theta)$ に対して，F が各 t とすべての $d \in D$ に対してベクトル (16.4.8) および，各 t に対するベクトル (16.4.9) を含むように拡大されるならば，E は閉じていて，(16.4.6) で定義されたベクトル列 $(L(\theta_1, d^*(x_t)), \cdots, L(\theta_h, d^*(x_t)))$ に対応する d^* は (16.4.1) で定義された平均危険 $\bar{r}(q, d)$ を最小にする．

(b)　幾何学的解釈

(16.4.10) の左辺を $\bar{r}(q, d^*)$ で表わせば，

(16.4.11) $$\bar{r}(q, d^*) = \min_{d \in D} \bar{r}(q, d) = \min_{r \in E} \sum_{i=1}^{h} r(\theta_i, d)q(\theta_i)$$

が成り立つ．よって，事前確率 $q(\theta_1), \cdots, q(\theta_h)$ を加重として，E の各点（各危険ベクトル）の座標を平均すれば，$\bar{r}(q, d^*)$ はこの加重平均の最小値に等しい．つまり，幾何学的に説明すると次のようになる．$q(\theta)$ が与えられたとき，R_h 空間内の E の点をすべて通る超平面の族

(16.4.12) $$\sum_{i=1}^{h} r_i q(\theta_i) = C$$

(C はパラメータである) を考える．このとき，最小の C を与える点 $r = (r_1, \cdots, r_h)$ が事前分布 $q(\theta)$ に対するベイズ解であり，この点に対する C の値が最小平均危険 $\bar{r}(q, d^*)$ となる．E が有界だから r のすべての座標は有限である．

(c)　すべての可能な事前分布に対する解集合

すべての可能な事前分布に対応する解集合について調べよう．

(16.4.13) $$(q(\theta_1), \cdots, q(\theta_h))$$

は h 個の点 $(1, 0, \cdots, 0), \cdots, (0, \cdots, 0, 1)$ で張られた R_h 内の単体 G の点であることに注意しておこう．点 (16.4.13) の座標は超平面 (16.4.12) の法線の傾きを示している．各可能な事前分布 $q(\theta)$ に対して最小の C を与える E の点 r が見い出されるならば，この最小な C の値の集合に対応する超平面の集合は，集合 E の下方からの境界包絡面を構成する．この包絡面に接する超平面上の E のすべての点 r の集合 E^* が包絡面自身の上にあることは明らかである．ベクトルの集合 E^* が，すべての可能な事前分布 $q(\theta)$ によって生成された決定問題のベイズ解全体の集合である．

例題　16.3 節の終わりで，ミニマックス解を得るために，簡単な数値例を与えた．ここで再度，その例題を考えよう．θ は $q(\theta_1) = \frac{1}{3}$，$q(\theta_2) = \frac{2}{3}$ なる事前分布を持つとする．ミニマックス解を求めるときの例において，決定関数 d_1，d_2，d_3，d_4 に対応する危険ベクトルは，それぞれ $(1, 3)$，$\left(\frac{7}{4}, \frac{5}{2}\right)$，$\left(\frac{13}{4}, \frac{5}{2}\right)$，$(4, 2)$ であった．この各ベクトルとベクトル $(q(\theta_1), q(\theta_2)) = \left(\frac{1}{3}, \frac{2}{3}\right)$ との内積 $\sum_i r(\theta_i, d)q(\theta_i)$ を計算する．対応する 4 つの内積はそれぞれ $\frac{28}{12}$, $\frac{27}{12}$, $\frac{33}{12}$, $\frac{32}{12}$ である．これは $q(\theta)$ に対して，それぞれ d_1，d_2，d_3，d_4 を用いたときの平均危険を示している．したがって，この（特定の）$q(\theta)$ に対する平均危険の最小値を達成する決定関数は d_2 である．前述した境界包絡線を 4 つの危険ベクトルから構成すれば，包絡線自身は d_1，d_2，d_4 を含んでいることがわかる．これは，d_1，d_2，d_4 のどれかが任意の可能な事前分布 $q(\theta)$ に対するベイズ解になり得ることを意味している．

16.5　拡張と一般化に関する注意

これまで非常に簡単な場合だけに限定して統計的決定論を議論してきた．すなわち，確

率変数 x は離散型で，パラメータ空間 Ω は有限個の点しか含まず，確率変数 x は一度だけしか観測されないとしてきた．（しかし，一般には，x は n 次元確率変数でもよい．たとえば，パラメータ θ を持つポアソン分布からの大きさ n の標本でもよい．しかし，パラメータ θ については有限個の値しかとらないと限定しておく）．このように比較的簡単な場合について，一般的なミニマックス解およびベイズ解を導き出した．

これらの結果の拡張および一般化に際しては種々の方向がある．まず x が密度関数 $f(x|\theta)$ を持つ絶対連続な確率変数で，パラメータ空間 Ω が有限個の場合への拡張が考えられる．次に，x は離散型でも連続型でもよいが，Ω が可付番無限個の点を含む場合，さらには，ユークリッド空間における点集合の場合への拡張もある．一般化の別の方向としては，逐次抜き取りを導入することも考えられる．しかしここでは，これらの拡張に対して一般論を試みることは，本書の範囲を超えていると思われる．この方面への興味を持たれる読者は，テキストとしては，Wald (1950), Blackwell と Girshick (1954), 文献としては Dvoretzky, Kiefer と Wolfowitz (1953 *a*, 1953 *b*), Karlin と Rubin (1956), Lehmann (1957) を参照されたい．

問　　題

16.1　x は 2 項分布

$$p(x|\theta) = \binom{2}{x}\theta^x(1-\theta)^{2-x}, \qquad x = 0, 1, 2$$

を持つ確率変数で，パラメータ θ の空間は 2 点 $\theta_1 = \frac{1}{100}$, $\theta_2 = \frac{1}{10}$ からなるとする．決定空間 A は 2 点 a_1, a_2 を含み，損失関数 $L(\theta_i, a_j)$ は次で定義されているとする．

	a_1	a_2
θ_1	1	4
θ_2	3	2

決定関数の空間 D は点 $d_1(x)$, $d_2(x)$, $d_3(x)$ からなっているとする．ただし，$a = 1, 2, 3$ に対して

$$d_\alpha(x) = \begin{cases} a_1, & x = 0, 1, \cdots, \alpha - 1 \\ a_2, & \text{その他}. \end{cases}$$

この決定問題のミニマックス解を与える決定関数を求めよ．

16.2（続き）　2 点 $\theta_1 = \frac{1}{100}$, $\theta_2 = \frac{1}{10}$ 上のすべての可能な事前分布に対するこの決定問題のベイズ解を与える決定関数を求めよ．

16.3　x は p.f.

$$p(x|\theta) = (1-\theta)\theta^{x-1}, \qquad x = 1, 2, \cdots$$

を持つ確率変数で，パラメータ空間 Ω は 2 点 $\theta_1 = \frac{1}{10}$, $\theta_2 = \frac{2}{10}$ からなるとする．決定空間 A は 2 つの要素 (a_1, a_2) を含み，損失関数 $L(\theta_i, a_j)$ は次のように定義されている．

	a_1	a_2
θ_1	1	3
θ_2	4	2

決定関数の空間 D は次の点からなるとする．

$$d_\alpha(x) = \begin{cases} a_1, & x = 1, \cdots, \alpha \\ a_2, & x = \alpha + 1, \alpha + 2, \cdots \end{cases}$$

ただし，$\alpha = 1, 2$．この決定問題のミニマックス解を与える決定関数を求めよ．

16.4（続き）　θ の空間は前問と同様に 2 点 $\theta_1 = \frac{1}{10}$, $\theta_2 = \frac{2}{10}$ だけからなると仮定して，θ のすべての可能な事前分布に対して決定問題のベイズ解を与える決定関数を求めよ．

16.5　x は $p(x|\theta_1)$, $p(x|\theta_2)$ のうち，どちらか一方を p.f. に持つ確率変数とする．決定空間 A は 2 つの要素 (a_1, a_2) からなり，損失関数 $L(\theta_i, a_j)$ は次の通りである．

	a_1	a_2
θ_1	0	c
θ_2	1	0

すべての可能な決定関数の中で，事前分布 $q(\theta)$ に対するベイズ解を与える決定関数 $d(x)$ は次のようになることを示せ．比に関する不等式

$$\frac{p(x|\theta_2)}{p(x|\theta_1)} \leqslant c\,\frac{q(\theta_1)}{q(\theta_2)}$$

を満たす x については，$d(x) = a_1$，その他の x に対しては $d(x) = a_2$ である．

第17章 時　系　列

17.1 概　説

理工学の分野では，あるプロセスを理想化して考えるとき，確率変数の族（確率過程）を生成する場合がよくある．すなわち，ある（時間）区間 T 内の各 t（時刻）の値に対して確率変数 x_t が対応するようなプロセスが考えられる．これは時間のランダム関数を生成していることにもなる．このような関数は**時系列**と呼ばれる．時系列の例として，回路の秒刻みの電圧，分刻みの工場内の騒音，時間刻みの潮の干満の高さ等がある．

他に，ある平滑傾向や時間関数を持つランダム変動の性質が見られる時系列もある．たとえば，このような時系列としては，ある町の 24 時間の気温，月毎の在庫価格，年毎の国家の雇用状態等が挙げられる．

時系列の研究および解析にあたって，種々の統計的および数学的方法が永年にわたって展開されてきた．この方法の多くは多少なりとも経験的な意味で時系列のランダム変動を“平滑化”することによって時系列に内在している平滑化傾向関数を推測するものであった．

しかし最近，時系列を確率過程として研究するような傾向が表われている．このようなアプローチは時系列が経験的な方法以外のやり方で生成され得ることも示している．

時系列を確率過程の立場で取り扱った文献がすでに幾冊かの著作とともに数多く発表されている．ここでは，2，3 の簡単で基本的な結果を含めて，時系列に関して手短かに紹介するだけにとどめよう．詳細な結果は多くの書物に見られる．特に，Bartlett (1955), Doob (1953), Grenander と Rosenblatt (1957), Wold (1938) には詳しく述べられている．

17.2 定常時系列

4.1 節では，確率過程は $t \in T$ の任意有限個の組，たとえば $t_1, \cdots, t_n$ に対する確率変数の組 $x_{t_1}, \cdots, x_{t_n}$ が結合確率分布関数を持つような確率変数の族 $\{x_t\,;\, t \in T\}$ として定義されていたことを思い出そう．

時系列においては，t を時刻，そして T は全時間軸（実軸）または時間軸の任意の区間あるいはその任意の部分集合とみなすことができる．このように，ある場合には T は時間軸上の等間隔の点列だけからなることもある．与えられた単位時間に対して，$T = \cdots, -1, 0, +1, \cdots$ ならば，T の有限ブロック（たとえば $1, \cdots, n$ とか $1, \cdots, M$ とか $n-h, n-h+1, \cdots, n-1$ とか $1, \cdots, 2k+1$ など）からなる標本が特に興味の中心となる．

任意の確率過程 $\{x_t\,;\, t \in T\}$（ただし，t は時刻を示す）はあまりにも一般的すぎて有益な議論はできない．ここでは，いわゆる**定常過程**あるいは**定常時系列**だけに限定しよう．2 つの重要な型が考えられる．

強定常時系列 $\{x_t\,;\, t \in T\}$ とは次の性質を持つ時系列である．すなわち，族 $\{x_t\,;\, t \in T\}$ からの確率変数の任意有限集合

(17.2.1)
$$(x_{t_1}, \cdots, x_{t_n})$$

が任意の h に対して

(17.2.2)
$$(x_{t_1+h}, \cdots, x_{t_n+h})$$

と同じ結合分布関数を持つ．（ほとんどの場合，1 次元確率変数を対象にしているが，確率変数 x_t は複素数値でも，k 次元確率変数でもよい．）このように，強定常時系列では異なる h の値に対する確率変数の組 (17.2.2) の結合分布はすべて等しく，時間差

(17.2.3)
$$t_2 - t_1, t_3 - t_2, \cdots, t_n - t_{n-1}$$

だけに依存する．

読者は容易に次を確かめることができよう．

17.2.1　$\{x_t\,;\, t \in T\}$ が $\mathscr{E}(|x_t|) < \infty$ なる強定常時系列ならば，すべての t について $\mathscr{E}(x_t) = c$（定数）である．

次に強定常時系列の具体例を与えよう．

例題　（実）時系列

(17.2.4) $$x_t=\sum_{p=1}^{r}a_p\cos(t\omega_p+u_p),\qquad t=\cdots,-1,0,+1,\cdots$$

と考える．ただし，$a_1,\cdots,a_r$ は実定数，$\omega_1,\cdots,\omega_r$ は区間 $[-\pi,+\pi]$ 上の定数で，$\omega_1<\cdots<\omega_r$ とする．$u_1,\cdots,u_r$ はおのおのが区間 $[-\pi,+\pi]$ 上の矩形分布に従う独立な確率変数である．このとき t の任意有限個の組 $t_1,\cdots,t_n$ に対して

(17.2.5) $$x_{t_\xi+h}=\sum_{p=1}^{r}a_p\cos(t_\xi\omega_p+v_p),\qquad \xi=1,\cdots,n$$

となる．ここに，$v_p\ (p=1,\cdots,r)$ は区間 $[h\omega_p\pm\pi]$ 上で一様分布する確率変数である．$(x_{t_1},\cdots,x_{t_n})$ の結合分布と $(x_{t_1+h},\cdots,x_{t_n+h})$ のそれが同一であることはすぐにわかる．よって，時系列 (17.2.4) は強定常である．

時系列 (17.2.4) に対応する複素数の場合は

(17.2.6) $$x_t=\sum_{p=1}^{r}a_pe^{i(t\omega_p+u_p)},\qquad t=\cdots,-1,0,+1,\cdots$$

である．もちろん，これも強定常である．

時系列の研究において，強定常を仮定せずに，次の (i)，(ii) の弱い条件の下で，いくつかの有益な結果が得られている． (i) すべての t に対して $\mathscr{E}(x_t)$ は定数（この定数を 0 にとってもよい）に等しい． (ii) (17.2.2) における分布はすべての h に対して同じ共分散行列を持つ．この 2 条件を満足する時系列 $\{x_t\,;\,t\in T\}$ は弱定常と呼ばれている．弱定常時系列においては，共分散行列は時間差 (17.2.3) だけに依存することを示している．したがって，x_{t+h} と x_t との共分散（$\mathscr{E}(x_t)=0$ とする）は h だけの関数，すなわち

(17.2.7) $$\mathscr{E}(x_{t+h}x_t)=\gamma_h.$$

h の関数として考えられた共分散 γ_h は時系列 $\{x_t\,;\,t\in T\}$ の**共分散関数**といわれる．また時間遅れ h を持つ**遅れ共分散**あるいは**自己共分散**と呼ばれることもある．x_{t+h} と x_t との相関係数は γ_h/γ_0（これを ρ_h で表わす）である．ρ_h を h の関数とみなしたとき，時系列の**自己相関関数**あるいは**時間相関関数**という．

x_t が複素数値ならば，その共分散関数 γ_h は

(17.2.8) $$\gamma_h=\mathscr{E}(x_{t+h}\bar{x}_t)$$

で定義される．ただし，$\bar{x}_t$ は x_t の共役複素数．

読者は容易に次を確かめることができよう．

17.2.2 有限な共分散行列を持つ強定常時系列はまた弱定常でもある．

ある時系列 $\{y_t\,;\,t\in T\}$ が次の型をしているとしよう．

(17.2.9) $$y_t=m_t+x_t.$$

ただし，m_t は各 t について定数，x_t は $\mathscr{E}(x_t)=0$ なる定常時系列 $\{x_t\,;\,t\in T\}$ の成分．このとき

(17.2.10) $$\mathscr{E}(y_t)=m_t$$

で，$\{y_t\,;\,t\in T\}$ の共分散関数は $\{x_t\,;\,t\in T\}$ のそれに等しい．すなわち，（実数値をとる時系列の場合）

(17.2.11) $$\gamma_h=\mathscr{E}(y_{t+h}-m_{t+h})(y_t-m_t)=\mathscr{E}(x_{t+h}x_t)$$

である．

時系列の問題の 1 つに，その時系列から抽出された有限な n 個の確率変数の観測に基づく m_t および γ_h の推定，あるいは，m_t および γ_h に関する仮説検定がある．もちろん，ここで観測される確率変数の個数 n が ∞ に近づく場合も考えられる．

17.3 定常時系列のスペクトル関数

(a) 特別な場合

定常時系列の最初の 2 つのモーメントに関する限り，時系列はその共分散関数 γ_h で表わされる．ここでは共分散関数それ自身が，いわゆる時系列のスペクトル分布関数で表現されることを示そう．

まず，(17.2.4) で与えられた特定の時系列の共分散関数と，そのスペクトル分布関数との関係を考える．すべての t に対して

(17.3.1) $$\mathscr{E}(x_t)=\sum_{p=1}^{r}a_p\mathscr{E}\cos(t\omega_p+u_p)=0$$

が成り立つことがすぐわかる．共分散関数 $\gamma_h=\mathscr{E}(x_{t+h}x_t)$ については

(17.3.2) $$\gamma_{t-s}=\sum_{p,q=1}^{r}a_pa_q\mathscr{E}\cos(s\omega_p+u_p)\cos(t\omega_q+u_q)$$

が成立する．しかし，$p\neq q$ のとき，$\mathscr{E}\cos(s\omega_p+u_p)\cos(t\omega_q+u_q)=0$．$p=q$ のとき

$$\cos(s\omega_p+u_p)\cos(t\omega_p+u_p)=\frac{1}{2}\cos((s+t)\omega_p+2u_p)+\frac{1}{2}\cos(s-t)\omega_p$$

だから

$$\mathscr{E}\cos(s\omega_p+u_p)\cos(t\omega_p+u_p)=\frac{1}{2}\cos(s-t)\omega_p$$

である．したがって，$t-s=h$ とおけば

$$\gamma_h=\frac{1}{2}\sum_{p=1}^{r}a_p^2\cos(h\omega_p)=\int_{-\pi}^{+\pi}\cos(h\omega)\,dF(\omega) \tag{17.3.3}$$

となる．ただし，$F(\omega)$ は

$$F(\omega)=\frac{1}{2}\sum_{\omega_p\leq\omega}a_p^2 \tag{17.3.4}$$

で定義された $[-\pi,+\pi]$ 上の ω の非減少階段関数である．$[-\pi,+\pi]$ の端点では，$F(-\pi)=0$，$F(\pi)=\gamma_0$ である．$F(\omega)$ は (17.3.3) の共分散関数 γ_h，$h=\cdots,-1,0,+1,\cdots$ に関する**スペクトル分布関数**と呼ばれる．(17.3.3) の右辺の表わし方は γ_h の**スペクトル表現**ともいわれる．

複素関数値時系列，すなわち (17.2.6) の場合には，共分散関数 γ_h は

$$\gamma_h=\frac{1}{2}\sum_{p=1}^{r}a_p^2e^{ih\omega_p}=\int_{-\pi}^{+\pi}e^{ih\omega}\,dF(\omega) \tag{17.3.5}$$

になることが簡単に計算される．ただし，$F(\omega)$ は (17.3.3) と同じスペクトル分布を持つ．

(b) 一 般 の 場 合

さて，与えられた時間単位に対して，$\mathscr{E}(x_t)\equiv0$ で共分散関数 γ_h を持つ任意の実定常時系列

$$x_t,\qquad t=\cdots,-1,0,+1,\cdots \tag{17.3.6}$$

を考えよう．この場合，$\gamma_{-h}=\gamma_h$ である．任意の整数 M に対して，$F_M(\omega)$ を次のように，$[-\pi,+\pi]$ 上の任意の ω に対して定義する．

$$\begin{aligned}F_M(\omega)&=\frac{1}{\pi M}\int_{-\pi}^{\omega}\mathscr{E}(x_1\cos\omega+\cdots+x_M\cos M\omega)^2\,d\omega \\ &=\frac{1}{\pi M}\int_{-\pi}^{\omega}\sum_{p,q=1}^{M}\gamma_{p-q}\cos p\omega\cos q\omega\,d\omega \\ &=\frac{1}{\pi M}\left\{\sum_{p\neq q=1}^{M}\gamma_{p-q}\left[\frac{\sin(p+q)\omega}{2(p+q)}+\frac{\sin(p-q)\omega}{2(p-q)}\right]\right. \\ &\qquad\left.+\gamma_0\sum_{p=1}^{M}\frac{\sin p\omega\cos q\omega}{2p}+\frac{\gamma_0M}{2}(\omega+\pi)\right\}.\end{aligned} \tag{17.3.7}$$

$F_M(\omega)$ は $[-\pi,+\pi]$ 上で非減少で，さらに $F_M(-\pi)=0$，$F_M(\pi)=\gamma_0$ である．$M\longrightarrow\infty$ ならば，$F_M(\omega)$ は，$F(-\pi)=0$，$F(\pi)=\gamma_0$ なる $[-\pi,+\pi]$ 上の非減少関数 $F(\omega)$ に収束する．(この極限関数 $F(\omega)$ の存在と一意性に関しては，**5.4.1** の証明を見られたい．) $dF_m(\omega)$，$dF(\omega)$ は $\omega=0$ のまわりで対称になる．

$\gamma_{-h}=\gamma_h$ だから，h が 0 と正整数の場合を考えれば十分である．$h=0,1,2,\cdots$ に対して

$$\gamma_h=\int_{-\pi}^{+\pi}\cos h\omega\,dF(\omega) \tag{17.3.8}$$

を示そう．整数列 $M_1,M_2,\cdots$ を考える．$h=0,1,\cdots,M_s$ に対して

$$\begin{aligned}\int_{-\pi}^{+\pi}\cos h\omega\,dF_{M_s}(\omega)&=\frac{1}{\pi M_s}\int_{-\pi}^{+\pi}\sum_{p,q=1}^{M_s}\gamma_{p-q}\cos p\omega\cos q\omega\cos h\omega\,d\omega \\ &=\frac{1}{\pi M_s}(A_h+B_h)\end{aligned}$$

が成り立つ．ただし

$$A_h=\int_{-\pi}^{+\pi}\sum_{p\neq q=1}^{M_s}\gamma_{p-q}\cos p\omega\cos q\omega\cos h\omega\,d\omega$$

$$B_h=\gamma_0\int_{-\pi}^{+\pi}\sum_{p=1}^{M_s}\cos^2p\omega\cos h\omega\,d\omega$$

である．さらに

$$\begin{aligned}\cos p\omega\cos q\omega\cos h\omega=\frac{1}{4}[&\cos(p+q+h)\omega+\cos(p+q-h)\omega \\ &+\cos(p-q+h)\omega+\cos(p-q-h)\omega]\end{aligned}$$

$$\cos^2p\omega\cos h\omega=\frac{1}{4}[\cos(2p+h)\omega+\cos(2p-h)\omega+2\cos h\omega]$$

$$\gamma_{p-q}=\gamma_{q-p}$$

を用いれば

$$A_h=\begin{cases}\pi\gamma_h(M_s-h), & h\neq0 \\ 0, & h=0\end{cases}\qquad B_h=\begin{cases}\pi\gamma_0\delta_h, & h\neq0 \\ \pi\gamma_0M_s, & h=0\end{cases}$$

を得る．ここに $\delta_h=\frac{1}{4}[1+(-1)^h]$ である．

したがって，$h=0,1,\cdots,M_s$ に対して

$$\int_{-\pi}^{+\pi}\cos h\omega\,dF_{M_s}(\omega)=\begin{cases}\left[\gamma_h\left(1-\dfrac{h}{M_s}\right)+\dfrac{\gamma_0\delta_h}{M_s}\right], & h\neq0 \\ \gamma_0, & h=0\end{cases} \tag{17.3.9}$$

が成立する．$M_s \longrightarrow \infty$ とすれば，任意の非負整数 h に対して (17.3.8) を得る．さて $\{F_M(\omega)\}$ の任意の収束部分列について同じ結果が得られる．したがって $\{F_M(\omega)\}$ 自身についても同じ結果を得る．$dF(\omega)$, $\cos h\omega$ はともに $\omega=0$ の回りで ω に関して対称だから，(17.3.8) は

(17.3.10) $$\gamma_h = 2\int_0^{\pi} \cos h\omega \, dF(\omega)$$

なる型に書かれる．

$F(\omega)$ を定常時系列 x_t のスペクトル分布関数という．$F(\omega)$ を γ_0 で割って基準化すれば，比 $F(\omega')/\gamma_0$ は，すべての経路 $\omega \leqslant \omega'$ により生成されたスペクトル質量の割合になる．$F(\omega)$ が導関数 $f(\omega)$ を持つ絶対連続関数ならば，$f(\omega)$ は時系列のスペクトル密度関数と呼ばれる．要約すると次が得られる．

17.3.1　x_t, $t=\cdots,-1,0,+1,\cdots$ は $\mathscr{E}(x_t)=0$ で，共分散関数 γ_h, $h=0,1,2,\cdots$ を持つ実定常時系列とする．このとき γ_h は (17.3.10) で与えられる．ただし，スペクトル分布関数 $F(\omega)$ は $\lim_{M\to\infty} F_M(\omega)$ で，$F_M(\omega)$ は (17.3.7) で与えられている．

共分散関数 $\gamma_h=\mathscr{E}(x_{t+h}\bar{x}_t)$, $h=\cdots,-1,0,+1,\cdots$ を持つ複素定常時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ に対応する **17.3.1** の定理は Herglotz (1911) により，次のように述べられている．

(17.3.11) $$\gamma_h = \int_{-\pi}^{+\pi} e^{ih\omega} \, dF(\omega)$$

ただし

(17.3.12) $$F(\omega) = \lim_{M\to\infty} F_M(\omega)$$

で

(17.3.13) $$F_M(\omega) = \frac{1}{2\pi M}\int_{-\pi}^{\omega} \mathscr{E}\left(\sum_{p=1}^{M} x_p e^{-ip\omega}\right)\left(\sum_{q=1}^{M} \bar{x}_q e^{iq\omega}\right) d\omega$$

である．被積分関数は実数で非負の値をとる．

複素数値の場合のこの定理の証明は **17.3.1** と同様である．読者に残しておこう．

実および複素定常時系列に対応する (17.3.10), (17.3.11) の結果は，物理的なプロセスで発生する多くの定常時系列問題に適用されるだろう．なぜなら，問題に応じて適当に時間単位を選べるからである．しかし，Bochner (1932) は (17.3.11) を h が単位時間内の整数倍の値をとる場合ではなくて，実（数）軸を動く場合に拡張している．それは次のようになる．複素定常時系列 $\{x_t\,;\,-\infty<t<\infty\}$ に対して，γ_h が $h=0$ で連続ならば

(17.3.14) $$\gamma_h = \int_{-\infty}^{\infty} e^{ih\omega} \, dF(\omega)$$

が成り立つ．実時系列の場合に対応する (17.3.10) の一般化は

(17.3.15) $$\gamma_h = 2\int_0^{\infty} \cos h\omega \, dF(\omega)$$

になる．

(c) 白　色　雑　音

実定常時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ が $[-\pi,+\pi]$ 上で定数スペクトル密度関数 $f(\omega)=\gamma_0/2\pi$ を持つならば，(17.3.8) より，すべての $h\neq 0$ に対して $\gamma_h=0$ になる．逆に，すべての $h\neq 0$ に対して，$\gamma_h=0$ とすれば

(17.3.16) $$\int_{-\pi}^{+\pi} \cos h\omega f(\omega)\, d\omega = \begin{cases} 0, & h \neq 0 \\ \gamma_0, & h=0 \end{cases}$$

である．$f(\omega)$ が $\omega=0$ のまわりで対称で，$b_0+b_1\cos\omega+b_2\cos 2\omega+\cdots$ にフーリエ展開できるならば，

(17.3.17) $$\int_{-\pi}^{+\pi} \cos h\omega f(\omega)\, d\omega = \int_{-\pi}^{+\pi} [b_0+b_1\cos\omega+b_2\cos 2\omega+\cdots]\cos h\omega\, d\omega, \quad h=0,1,\cdots$$

である．しかし，(17.3.17) は

(17.3.18) $$\int_{-\pi}^{+\pi} \cos h\omega f(\omega)\, d\omega = b_h\int_{-\pi}^{+\pi} \cos^2 h\omega\, d\omega$$

となる．(17.3.16) で述べているように，左辺は $h\neq 0$ のとき 0．よって，$h\neq 0$ ならば，$b_h=0$．ゆえに，$f(\omega)=b_0$ である．また

$$\int_{-\pi}^{+\pi} f(\omega)\, d\omega = \gamma_0$$

より，$b_0=\gamma_0/2\pi$ になる．したがって，$[-\pi,+\pi]$ 上で

(17.3.19) $$f(\omega) = \frac{\gamma_0}{2\pi}$$

となる．よって，次を得る．

17.3.3　実定常時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ が $[-\pi,+\pi]$ 上で $b_0+b_1\cos\omega+b_2\cos 2\omega+\cdots$ にフーリエ展開される対称なスペクトル密度関数 $f(\omega)$ を持つとする．すべての $h\neq 0$ に対する γ_h が 0 になるための必要十分条件は $f(\omega)$ が $[-\pi,+\pi]$ 上で定数になることである．このとき $f(\omega)=\gamma_0/2\pi$ である．

すべての成分の対 x_t, $x_{t'}$, $t\neq t'$ が 0 の相関を持つ定常時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ は白色雑音と呼ばれる．すなわち，**17.3.3** は時系列がスペクトル密度関数のある特定のクラスに対して白色雑音になるための必要十分条件を述べている．

17.4　定常時系列の平均関数および共分散関数の推定

(a) 平 均 の 推 定

いま，x_t, $t=\cdots,-1,0,+1,\cdots$ を $\mathscr{E}(x_t)=\mu$ なる実定常時系列とする．このとき，有限個の標本 $(x_1,\cdots,x_n)$ から μ を推定したい．平均 $\bar{x}=\frac{1}{n}(x_1+\cdots+x_n)$ は μ の不偏推定量である．なぜなら

(17.4.1)
$$\mathscr{E}(\bar{x})=\mu$$

である．$\bar{x}$ の分散は

(17.4.2)
$$\sigma^2(\bar{x})=\frac{1}{n}\Big[\gamma_0+2\sum_{\xi=1}^{n-1}\Big(1-\frac{\xi}{n}\Big)\gamma_\xi\Big]$$

で与えられる．

$F(\omega)$ が導関数 $f(\omega)$ を持ち，$f(\omega)$ が $\omega=0$ のまわりで対称で，$\omega=0$ で連続で，さらに $[-\pi,+\pi]$ 上でフーリエ級数

(17.4.3)
$$\frac{1}{2\pi}(\gamma_0+2\gamma_1\cos\omega+2\gamma_2\cos 2\omega+\cdots)$$

に展開できるならば，$n\longrightarrow\infty$ のとき (17.4.2) の [ ] 内の量は極限値 $2\pi f(0)$ を持つことが確かめられる．したがって，これらの条件の下では十分に大きな n に対して

(17.4.4)
$$\sigma^2(\bar{x})\cong\frac{2\pi f(0)}{n}$$

が成り立つ．

(b) 分 散 の 推 定

γ_h の推定量を

(17.4.5)
$$c_h=\frac{1}{n-h}\sum_{\xi=1}^{n-h}(x_\xi-\bar{x})(x_{\xi+h}-\bar{x})$$

とすれば，代数的な演算により

(17.4.6)
$$\mathscr{E}(c_h)=\gamma_h-\sigma^2(\bar{x})+\frac{2}{n}R$$

となる．ただし

(17.4.7)
$$R=\Big[\sum_{\xi=1}^{n-1}\Big(1-\frac{\xi}{n}\Big)\gamma_\xi-\sum_{\xi=1}^{n-h-1}\Big(1-\frac{\xi}{n-h}\Big)(\gamma_\xi+\gamma_{\xi+h})-\sum_{\xi=1}^{h}\gamma_\xi\Big].$$

$n\longrightarrow\infty$ のとき $R\longrightarrow 0$ である．よって n が十分に大きいとき，γ_h の推定量としての c_h の漸近的な偏りは近似的に $-\sigma^2(\bar{x})$ になる．スペクトル分布が $\omega=0$ で連続な密度関数 $f(\omega)$ を持てば，十分大きな n に対して

(17.4.8)
$$\mathscr{E}(c_h)\cong\gamma_h-\frac{2\pi f(0)}{n}$$

が成立する．

μ が既知のとき (17.4.5) で $\bar{x}$ のかわりに μ を代入すれば，この共分散の平均値は γ_h に一致することに注意しよう．

c_h の分散を求める問題を考えるときは，時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ の4次のモーメントが有限であるという仮定とその構造に関する若干の仮定を設ける必要があるだろう．しかし x_t, $t=\cdots,-1,0,+1,\cdots$ が強定常ならば c_h の分散に関する近似式が得られる．まず，時系列 $z_t=(x_t-\mu)(x_{t+h}-\mu)$, $t=\cdots,-1,0,+1,\cdots$ を考えよう．x_t が強定常のとき z_t も強定常になる．この時系列の平均は γ_h である．その共分散関数が存在するとして，これを $\gamma^*_{h,g}$ で表わそう．すなわち

(17.4.9)
$$\gamma^*_{h,g}=\mathscr{E}(z_tz_{t+g})$$

とする．

さて，n が十分に大きいとき，c_h の分散は次の c^*_h の分散で近似されることが示される．ただし

(17.4.10)
$$c^*_h=\frac{1}{n-h}\sum_{\xi=1}^{n-h}x_\xi x_{\xi+h}.$$

c^*_h の分散は $\sigma^2(\bar{x})$ と同じ構造を持ち

$$\sigma^2(c_h^*) = \frac{1}{n-h}\left[\gamma_{h,0}^* + 2\sum_{\xi=1}^{n-h-1}\left(1-\frac{\xi}{n-h}\right)\gamma_{h,\xi}^*\right] \tag{17.4.11}$$

で与えられる．

時系列 z_t がスペクトル密度関数を持てば，この級数は収束して十分大きな n に対して

$$\sigma^2(c_h^*) \cong \frac{\gamma_{h,0}^*\left(3-\frac{1}{\pi}\right)}{n} \tag{17.4.12}$$

が成り立つ．ここで

$$\gamma_{h,0}^* = \mathscr{E}[(x_t-\mu)(x_{t+h}-\mu)]^2$$

に注意しよう．

17.5 スペクトル分布の推定

定常時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ は平均 μ とスペクトル密度関数 $f(\omega)$ を持つとする．この節では μ を既知として，時系列からの標本 $(x_1,\cdots,x_n)$ に基づいて $f(\omega)$ を推定する問題を考察しよう．

(17.3.7) で定義された $F_M(\omega)$ の $M\longrightarrow\infty$ のときの極限として $F(\omega)$ が定義されたことを考慮すれば

$$f_n(\omega) = \frac{1}{n\pi}(y_1\cos\omega + y_2\cos 2\omega + \cdots + y_n\cos n\omega)^2 \tag{17.5.1}$$

で定義される $f_n(\omega)$ が $(x_1,\cdots,x_n)$ に基づく $f(\omega)$ の推定量になっていると予想される（ただし，$y_t = x_t-\mu$）．この平均値をとり簡単な 3 角級数の和をとると

$$\mathscr{E}f_n(\omega) = \frac{1}{2\pi}\left\{\gamma_0\left[1+\frac{\sin n\omega\cos(n+1)\omega}{n\sin\omega}\right] + 2\sum_{\xi=1}^{n-1}\gamma_\xi\left[\left(1-\frac{\xi}{n}\right)\cos h\omega + \frac{\sin(n-\xi)\omega\cos(n+1)\omega}{n\sin\omega}\right]\right\} \tag{17.5.2}$$

になる．$n\longrightarrow\infty$ とし，(17.4.3) で与えられる $f(\omega)$ のフーリエ展開を用いれば，$\omega=0$ 以外の $[-\pi,+\pi]$ 上の任意の ω に対して

$$\lim_{n\to\infty}\mathscr{E}f_n(\omega) = f(\omega) \quad \text{で，} \quad \lim_{n\to\infty}\mathscr{E}f_n(0) = 2f(0)$$

となる．このように $\omega=0$ を除外すれば，$f_n(\omega)$ は $[-\pi,+\pi]$ 上の各 ω について $f(\omega)$ の漸近的な不偏推定量になっている．他方，比較的弱い条件の下で，$f_n(\omega)$ の分散は $n\longrightarrow\infty$ のとき，0 に収束しないことが示される．特に $(x_1,\cdots,x_n)$ が n 次元正規確率変数ならば，$f_n(\omega)$ は，すべての n について定数倍をのぞけば，自由度 1 のカイ 2 乗分布を持つ．

よって $f(\omega)$ の推定量 $f_n(\omega)$ は一致推定量ではない．したがって $f(\omega)$ に対する別の推定量が考えられねばならない．$f(\omega)$ に対する種々の一致推定量が Bartlett (1950), Grenander と Rosenblatt (1953), Jenkins と Priestly (1957), Parzen (1957), Tukey (1949 *a*) 等によって提案され，取り扱われている．これらの多くの方法は，区間 $[-\pi,+\pi]$ 上の固定された点 ω における $f(\omega)$ の推定問題を議論している．しかし，トゥキーの方法は $[-\pi,+\pi]$ の部分区間内のスペクトル質量を推定する問題を考察している．この分野に興味を持たれた読者は Bartlett (1950), Grenander と Rosenblatt (1957) の著作を参照されたい．

ここでは標本 $(x_1,\cdots,x_n)$ が抽出された時系列に対してフーリエ展開 (17.4.3) されるスペクトル密度関数 $f(\omega)$ が存在すると仮定しよう．さて与えられた正整数 m に対して，標本 $(x_1,\cdots,x_n)$ から区間

$$I_p = \left(\frac{p\pi}{m}-\frac{\pi}{2m},\ \frac{p\pi}{m}+\frac{\pi}{2m}\right], \qquad p=0,\pm1,\pm2,\cdots,\pm m \tag{17.5.3}$$

に含まれるスペクトル質量

$$\Delta_p = F\left(\frac{p\pi}{m}+\frac{\pi}{2m}\right) - F\left(\frac{p\pi}{m}-\frac{\pi}{2m}\right) \tag{17.5.4}$$

を推定する方法を考えよう．ここに $F(\omega)$ はスペクトル分布関数である．

数 m は，スペクトルを正確に調べたいために必要以上の多くのスペクトル質量を推定する場合に生じる不確定要素を均一化するように選べばよい．これは大ざっぱにいえば，時系列の共分散関数が具体的に表現されるように m を共分散の数に選べばよいことを意味している．

スペクトル質量 Δ_p は $f(\omega)$ を用いると

$$\Delta_p = \int_{\frac{p\pi}{m}-\frac{\pi}{2m}}^{\frac{p\pi}{m}+\frac{\pi}{2m}} f(\omega)\,d\omega \tag{17.5.5}$$

のように表わされる．

$f(\omega)$ に対する (17.4.3) のフーリエ展開を用いて積分すると

(17.5.6)
$$\Delta_p = \frac{\gamma_0}{2m} + \frac{1}{m}\sum_{h=1}^{\infty}\gamma_h\left(\frac{\sin(h\pi/2m)}{h\pi/2m}\right)\cos\frac{hp\pi}{m}$$

が得られる.

さて時系列において情報が共分散 $\gamma_0, \gamma_1, \cdots, \gamma_m$ によって具体的に表現されていると考えれば, スペクトル質量 Δ_p, $p = 0, \pm 1, \cdots, \pm m$ は (17.5.6) の無限級数を $h = m$ で打ち切って近似される. すなわち Δ_p の近似スペクトル質量

(17.5.7)
$$\tilde{\Delta}_p = \frac{\gamma_0}{2m} + \frac{1}{m}\sum_{h=1}^{m}\gamma_h\left(\frac{\sin(h\pi/2m)}{h\pi/2m}\right)\cos\frac{hp\pi}{m}$$

が得られる. ただし, $p = 0, \pm 1, \cdots, \pm m$.

この時系列の μ が既知で, 標本 $(x_1, \cdots, x_n)$ が抽出されたならば, $\tilde{\Delta}_p$ の推定量は $\tilde{\Delta}_{pn}$ になる. ただし,

(17.5.8)
$$\tilde{\Delta}_{pn} = \frac{c_0}{2m} + \frac{1}{m}\sum_{h=1}^{m}c_h\left(\frac{\sin(h\pi/2m)}{h\pi/2m}\right)\cos\frac{hp\pi}{m}$$

で

$$c_h = \frac{1}{n-h}\sum_{\xi=1}^{n-h}y_\xi y_{\xi+h}, \qquad h = 0, 1, \cdots, n-1, \quad m \leqslant n-1, \quad y_\xi = x_\xi - \mu$$

である. $\mathscr{E}(c_h) = \gamma_h$, $h = 0, 1, \cdots$ より明らかに

(17.5.9)
$$\mathscr{E}(\tilde{\Delta}_{pn}) = \tilde{\Delta}_p.$$

すなわち $\tilde{\Delta}_{pn}$ は $\tilde{\Delta}_p$ の不偏推定量である.

ここでスペクトル質量の推定量 $\tilde{\Delta}_{pn}$ は Tukey (1949 *a*) が提唱した推定量と微妙に異なっていることを注意しておこう. $\dfrac{\sin(h\pi/2m)}{h\pi/2m}$ をその近似値 $(0.46\cos(h\pi/m)+0.54)$ で置き換え, 端点では (17.5.8) における和の最後の項を $\dfrac{1}{2}$ 倍すれば Δ_p に対するトゥキーの推定量が得られる.

$\tilde{\Delta}_{pn}$ は区間 $p\pi/m \pm \pi/2m$ におけるスペクトル質量の推定量だから, 区間 I_p 上の平均スペクトル密度 $\bar{f}_p(\omega)$ の不偏推定量 $\mathscr{E}^{-1}(\bar{f}_p(\omega))$ はスペクトル質量の推定量 $\tilde{\Delta}_{pn}$ を区間の長さ π/m で割れば得られる. すなわち

(17.5.10)
$$\mathscr{E}^{-1}(\bar{f}_p(\omega)) = \frac{m}{\pi}\tilde{\Delta}_{pn}.$$

(17.5.10) における推定量 $\dfrac{m}{\pi}\tilde{\Delta}_{pn}$ の分散は, 十分大きな n と m (ただし $m = 0(n)$) に対しては

(17.5.11)
$$\sigma^2\left(\frac{m}{\pi}\tilde{\Delta}_{pn}\right) \cong \frac{2}{n}\bar{f}_p^2(\omega)\sum_{h=-m}^{m}\left(\frac{\sin(h\pi/2m)}{h\pi/2m}\right)^2\cos^2\frac{hp\pi}{m}$$

となる.

時系列の平均 μ が未知ならば, c_h の定義において, μ のかわりに標本平均 $\bar{x}$ を代用する. こうして定義された c_h, $h = 0, 1, \cdots, m$ を (17.5.8) に代入するとスペクトル質量 $\tilde{\Delta}_{pn}$ の推定量が得られる. このように c_h の定義をかえても (17.5.9), (17.5.11) はともに十分大きな n に対して成立する.

17.6　パラメトリック時系列の統計的検定

定常時系列に関する種々の統計的推測問題を詳細に取り扱うことは本書の範囲を超えるので, ここでは, いくつかの古典的なパラメトリック時系列の場合にだけ限定して推定および統計的検定を論じよう. この方面に深く興味を持たれる読者は Bartlett (1955), Grenander と Rosenblatt (1957) を参照されたい.

(a) 階　差　法

時系列 x_t, $t = \cdots, -1, 0, +1, \cdots$ は

(17.6.1)
$$x_t = \sum_{p=0}^{k}\beta_p t^p + \varepsilon_t = \mu_t + \varepsilon_t$$

なる型をしているとする. ただし $\beta_0, \beta_1, \cdots, \beta_k$ は未知で, ε_t は分散 σ^2 を持つ白色雑音である.

k が既知ならば標本 $(x_1, \cdots, x_n)$, $n \geqslant k+1$ に対して, β の最小分散推定量は最小2乗法によって得られ, この推定量の分散は 10.3 (c) 節の方法で計算される. いま, 各 t^p を $p\ (=0, 1, \cdots, k)$ 次の多項式 $g_p(t)$ に置き換えても $g_0(t), g_1(t), \cdots, g_k(t)$ が関数的に独立ならば同様な方法が適用でき, 最小分散推定量およびその分散が求められる.

しかし (17.6.1) において k が未知の場合, k の推定量を求める方法の 1 つとして階差法といわれる半ば経験的な方法が適用できる. これは次のようにする. $y_{h,t}$ を時系列 x_t の h 次の前向きの差分で定義される時系列とする. すなわち

(17.6.2)
$$y_{h,t}=\Delta^h x_t=\Delta^h\mu_t+\Delta^h\varepsilon_t.$$
ただし

(17.6.3)
$$\Delta^h\varepsilon_t=\varepsilon_{t+h}-\binom{h}{1}\varepsilon_{t+h-1}+\binom{h}{2}\varepsilon_{t+h-2}-\cdots+(-1)^h\varepsilon_t$$

で $\Delta^h\mu_t$ も同様である．さて標本列 $(x_1,\cdots,x_{n+h})$，$h=1,2,\cdots$ を考え，比

(17.6.4)
$$Q_h=\sum_{\xi=1}^{n}y_{h,\xi}^2\Big/\left[n\binom{2h}{h}\right],\qquad h=1,2,\cdots$$

をつくる．

Q_h の平均値をとると

(17.6.5)
$$\mathscr{E}Q_h=\mathscr{E}\sum_{\xi=1}^{n}(\Delta^h\mu_\xi+\Delta^h\varepsilon_\xi)^2\Big/\left[n\binom{2h}{h}\right]$$

を得る．ε_ξ は白色雑音で μ_ξ は t の非ランダム関数だから

(17.6.6)
$$\mathscr{E}(\Delta^h\mu_\xi+\Delta^h\varepsilon_\xi)^2=(\Delta^h\mu_\xi)^2+\mathscr{E}(\Delta^h\varepsilon_\xi)^2.$$

しかも

(17.6.7)
$$\mathscr{E}(\Delta^h\varepsilon_\xi)^2=\sigma^2\sum_{p=0}^{h}\binom{h}{p}^2=\binom{2h}{h}\sigma^2.$$

よって (17.6.7) を (17.6.6) に代入し，さらに (17.6.5) に代入すると

(17.6.8)
$$\mathscr{E}Q_h=R_h+\sigma^2,\qquad h=1,2,\cdots$$

を得る．ただし

$$R_h=\frac{1}{n}\sum_{\xi=1}^{n}(\Delta^h\mu_\xi)^2\Big/\binom{2h}{h}$$

である．

R_h は非負で，$h\geqslant k+1$ のとき 0 になる．よって $h\geqslant k+1$ ならば，$\mathscr{E}Q_h=\sigma^2$ である．

このように実際問題では Q_h，$h=1,2,\cdots$ を逐次計算していって k の推定値として，Q_h が定数になる（ただし“微小な”変動を許す）ような h の値を選べばよい．このときの Q_h 自身がまた σ^2 の推定量にもなっている．階差法，その変形された方法およびその応用等に深く興味を持たれる読者は Tintner (1940), Quenouille (1948) を参照されるとよい．

(b)　既知の係数と周期を持つ 3 角時系列

時系列 x_t，$t=\cdots,-1,0+1,\cdots$ は次のように周期的で，しかもパラメータを含む型に表現されているとする．すなわち

(17.6.9)
$$x_\xi=\sum_{p=1}^{k}[a_p\cos\omega_p\xi+b_p\sin\omega_p\xi]+\varepsilon_\xi.$$

ここに ε_ξ は既知の分散 σ^2 を持つ白色雑音，a_p，b_p，ω_p は既知の［実］定数で $0\leqslant\omega_p\leqslant\pi$ とする．

さてこの時系列からの標本 $(x_1,\cdots,x_n)$ を考えよう．(17.6.9) に $\cos\omega\xi$ 倍 $(0\leqslant\omega\leqslant\pi)$ して，ξ で和をとり，n で割れば

(17.6.10)
$$\alpha(\omega)=\sum_{p=1}^{k}[a_pu_p(\omega)+b_pv_p(\omega)]+\delta(\omega)$$

が得られる．ここに

(17.6.11)
$$\alpha(\omega)=\frac{1}{n}\sum_{\xi=1}^{n}x_\xi\cos\omega\xi$$
$$u_p(\omega)=\frac{1}{n}\sum_{\xi=1}^{n}\cos\omega_p\xi\cos\omega\xi$$
$$v_p(\omega)=\frac{1}{n}\sum_{\xi=1}^{n}\sin\omega_p\xi\cos\omega\xi$$
$$\delta(\omega)=\frac{1}{n}\sum_{\xi=1}^{n}\varepsilon_\xi\cos\omega\xi.$$

同様に，(17.6.9) に $\sin\omega\xi$，$0\leqslant\omega\leqslant\pi$ を掛けて，ξ で和をとり，n で割ると

(17.6.12)
$$\alpha'(\omega)=\sum_{p=1}^{k}[a_pu_p'(\omega)+b_pv_p'(\omega)]+\delta'(\omega)$$

となる．ここで $\alpha'(\omega)$，$u_p'(\omega)$，$v_p'(\omega)$，$\delta'(\omega)$ は，それぞれ $\alpha(\omega)$，$u_p(\omega)$，$v_p(\omega)$，$\delta(\omega)$ において $\sin\omega\xi$ を $\cos\omega\xi$ に置き換えたものである．

$u_p(\omega)$，$v_p(\omega)$，$u_p'(\omega)$，$v_p'(\omega)$ は次のように書ける．

$$u_p(\omega)=\frac{1}{2n}\sum_{\xi=1}^{n}[\cos(\omega_p+\omega)\xi+\cos(\omega_p-\omega)\xi]$$

(17.6.13)
$$v_p(\omega)=\frac{1}{2n}\sum_{\xi=1}^{n}[\sin(\omega_p+\omega)\xi+\sin(\omega_p-\omega)\xi]$$

$$u'_p(\omega) = \frac{1}{2n}\sum_{\xi=1}^{n}[\sin(\omega_p+\omega)\xi - \sin(\omega_p-\omega)\xi]$$

$$v'_p(\omega) = \frac{1}{2n}\sum_{\xi=1}^{n}[-\cos(\omega_p+\omega)\xi - \cos(\omega_p-\omega)\xi].$$

よって $n \longrightarrow \infty$ とすれば

(17.6.14)
$$\mathscr{E}[\alpha(\omega)] \longrightarrow \begin{cases} 0, & \omega \neq \omega_p, \quad p=1,\cdots,k \\ \dfrac{a_p}{2}, & \omega = \omega_p \end{cases}$$

が成立する．同様に $n \longrightarrow \infty$ のとき

(17.6.15)
$$\mathscr{E}[\alpha'(\omega)] \longrightarrow \begin{cases} 0, & \omega \neq \omega_p, \quad p=1,\cdots,k \\ -\dfrac{b_p}{2}, & \omega = \omega_p \end{cases}$$

である．このように $2\pi/\omega_p$ が時系列 $x_t,\ t=\cdots,-1,0,+1,\cdots$ の真の周期ならば，大標本では $\alpha(\omega_p)$ は $\frac{a_p}{2}$ に，$\alpha'(\omega_p)$ は $-\frac{b_p}{2}$ に近づく傾向にある．

いま，時系列が与えられていて，その周期 $2\pi/\omega_p,\ p=1,\cdots,k$ が既知で，係数 a_p, b_p, $p=1,\cdots,k$ が未知ならば，10.3 (c) 節より最小 2 乗法によって，これらの係数の最小分散線形推定量を求めることができる．

(c)　未知のパラメータを含む 3 角時系列——ペリオドグラム解析

Schuster (1898) は時系列が (17.6.9) の型をしていて，a_p, b_p, ω_p, k がいずれも未知である場合について，可能な周期を見いだす方法を提案した．Fisher (1929) はあらかじめ予想された周期の有意性検定の方法を提唱している．ここではペリオドグラム解析として知られているこの方法を考察しよう．

前にも述べたように関数 $\alpha(\omega)$, $\alpha'(\omega)$ の平均値の動向を考慮すると，ω の関数として考えられた $\alpha(\omega)$, $\alpha'(\omega)$ は真の周期が存在すれば，その周期を求めるのに非常に有用であろうと思われる．真の周期がまったく存在しないと仮定しよう．すなわち，すべての a_p, b_p は 0 とする．このとき x_t は白色雑音になる．n が奇数，たとえば $n=2k+1$ のとき

(17.6.16)
$$\omega_p = \frac{2\pi p}{2k+1}, \qquad p=1,\cdots,k$$

とする．さて $\alpha(\omega_p)$, $\alpha'(\omega_p)$, $p=1,\cdots,k$ はすべて x_t の線形関数である場合を考えよう．$\alpha(\omega_p)$ と $\alpha(\omega_q)$, $p \neq q$ の分散は

(17.6.17)
$$\operatorname{cov}(\alpha(\omega_p),\alpha(\omega_q)) = \frac{\sigma^2}{(2k+1)^2}\sum_{\xi=1}^{2k+1}\cos\omega_p\xi\cos\omega_q\xi$$
$$= \frac{\sigma^2}{2(2k+1)^2}\sum_{\xi=1}^{2k+1}[\cos(\omega_p+\omega_q)\xi + \cos(\omega_p-\omega_q)\xi].$$

しかし

$$\sum_{\xi=1}^{2k+1}\cos(\omega_p\pm\omega_q)\xi = \left\{\frac{e^{i(\omega_p\pm\omega_q)}[1-e^{i(\omega_p\pm\omega_q)(2k+1)}]}{1-e^{i(\omega_p\pm\omega_q)}}\right\} \text{ の実部}$$

である．$(\omega_p\pm\omega_q)(2k+1)$ は 2π の整数倍であるから [　] 内の値は 0 になる．ゆえに

$$\operatorname{cov}(\alpha(\omega_p),\alpha(\omega_q)) = 0, \qquad p \neq q.$$

同様にすると

$$\operatorname{cov}(\alpha'(\omega_p),\alpha'(\omega_q)) = 0, \qquad p \neq q$$

で

$$\operatorname{cov}(\alpha'(\omega_p),\alpha(\omega_q)) = 0, \qquad p,q=1,\cdots,k.$$

さらに，$p=q$ のとき (17.6.17) は $\alpha(\omega_p)$ の分散は

$$\sigma^2(\alpha(\omega_p)) = \frac{\sigma^2}{2(2k+1)}$$

になる．同様に

$$\sigma^2(\alpha'(\omega_p)) = \frac{\sigma^2}{2(2k+1)}.$$

要約すると次の結果を得る．

17.6.1　$x_1,\cdots,x_{2k+1}$ は平均 0，分散 σ^2 を持つ独立な確率変数である．このとき，ω_p を (17.6.16) で与えると，$\alpha(\omega_p)$, $\alpha'(\omega_p)$, $p=1,2,\cdots,k$ は平均と共分散が 0 で，分散が $\sigma^2/(4k+2)$ の確率変数である．

ここで標本 $(x_1,\cdots,x_{2k+1})$ が長さ $2\pi/\omega_p,\ p=1,\cdots,2k+1$ の可能な周期を持つかどうかを調べてみよう．$2\pi/\omega_p$ が真の周期ならばすでに $\alpha(\omega_p)$, $\alpha'(\omega_p)$ の性質で述べたように，十分大きな $n\ (=2k+1)$ に対して $\alpha(\omega_p)$ と $\alpha'(\omega_p)$ はともに 0 からかけ離れた値を持つ．これはいい換えれば，$2\pi/\omega_p$ が真の周期であるとき，$\alpha^2(\omega_p)+\alpha'^2(\omega_p)$ が有意に大きな正の値に近づくことを意味している．しかし標本 $(x_1,\cdots,x_{2k+1})$ が与えら

れたとき $\alpha^2(\omega_p)+\alpha'^2(\omega_p)$ は各 ω_p, $p=1,\cdots,k$ に対して 1 つの値を持つ．それで $\{x_t\}$ が白色雑音であるという仮定の下で，$\alpha^2(\omega_p)+\alpha'^2(\omega_p)$, $p=1,\cdots,k$ のうち最も（または m 番目に）大きい値が有意に大きいかどうかを検定する（したがって検定する方法も）必要があるだろう．もちろん，一般には σ^2 が未知であるために複雑になるだろう．次の量は $(x_1,\cdots,x_{2k+1})$ のペリオドグラムと呼ばれる．

(17.6.18)
$$I_{2k+1}(\omega)=\frac{(2k+1)}{2\pi}[\alpha^2(\omega)+\alpha'^2(\omega)].$$

有意性検定の問題を取り扱うために，さらに $\{x_t\}$ が正規白色雑音であることを仮定する．すなわち $x_1,\cdots,x_{2k+1}$ は独立で，すべて分布 $N(0,\sigma^2)$ を持つとする．このとき

$$\frac{\sqrt{2(2k+1)}}{\sigma}\alpha(\omega_p),\qquad \frac{\sqrt{2(2k+1)}}{\sigma}\alpha'(\omega_p),\qquad p=1,\cdots,k$$

は $2k$ 個の独立確率変数で正規分布 $N(0,1)$ を持つ．いま

(17.6.19)
$$u_p=\frac{2k+1}{\sigma^2}[\alpha^2(\omega_p)+\alpha'^2(\omega_p)]=\frac{2\pi}{\sigma^2}I_{2k+1}(\omega_p),\qquad p=1,\cdots,k$$

とすると，$u_1,\cdots,u_k$ は $u_p\geqq 0$, $p=1,\cdots,k$ 上で確率素分

(17.6.20)
$$e^{-(u_1+\cdots+u_k)}\,du_1\cdots du_k.$$

その他では 0 を持つ独立な確率変数になる．$\alpha^2(\omega_p)+\alpha'^2(\omega_p)$（または $I_{2k+1}(\omega_p)$）のうち最大（あるいは m 番目に大きい）値が有意に大きいかどうかという問題は $u_1,\cdots,u_k$ の最大値が有意に大きいかどうかを検定する問題に帰着する．σ^2 は未知だから，この検定は，比

(17.6.21)
$$\frac{u_p}{u_1+\cdots+u_k},\qquad p=1,\cdots,k$$

の最大値（あるいは m 番目に大きい値）が有意に大きいかどうかを検定することである．この比は σ^2 に無関係であることに注意しよう．

v_r $(r=1,\cdots,k)$ を $u_1,\cdots,u_k$ のうち r 番目に小さい u とする．すなわち $v_1,\cdots,v_k$, $0<v_1<v_2<\cdots<v_k<+\infty$ は $u_1,\cdots,u_k$ の順序統計量で，その確率素分は

(17.6.22)
$$k!\,e^{-(v_1+\cdots+v_k)}\,dv_1\cdots dv_k$$

となる．ここで

(17.6.23)
$$g=\frac{v_k}{v_1+\cdots+v_k}$$

とすれば g は (17.6.21) における比の最大値と恒等的に同じ確率変数である．まず以下の簡単ではあるが有効な方法によって g の h 次のモーメントを計算してから g の分

布関数を求めよう．

g の h 次のモーメントは

(17.6.24)
$$\mathscr{E}(g^h)=\int_{-\infty}^0\cdots\int_{-\infty}^0\mathscr{E}\{e^{(v_1+\cdots+v_k)\varphi}v_k^h\}\,dz_1\cdots dz_h$$

で与えられることに注意しよう．ただし $\varphi=(z_1+\cdots+z_h)$．しかし $v_1,\cdots,v_{k-1}$ で積分すると

(17.6.25)
$$\mathscr{E}\{e^{(v_1+\cdots+v_k)\varphi}v_k^h\}=k(1-\varphi)^{-(k-1)}\int_0^\infty(1-e^{-v_k(1-\varphi)})^{k-1}e^{-v_k(1-\varphi)}v_k^h\,dv_k.$$

これを (17.6.24) に代入して $(1-e^{-v_k(1-\varphi)})^{k-1}$ を展開して v_k で積分する（全体として $z_1,\cdots,z_h$ で積分することになる）と

(17.6.26)
$$\mathscr{E}(g^h)=k(k-1)\sum_{p=0}^{k-1}(-1)^p\binom{k-1}{p}\frac{\Gamma(h+1)\Gamma(k-1)}{(p+1)^{h+1}\Gamma(h+k)}$$

が得られる．しかし

(17.6.27)
$$\frac{\Gamma(h+1)\Gamma(k-1)}{\Gamma(h+k)}=\int_0^1 w^h(1-w)^{k-2}\,dw.$$

よって

(17.6.28)
$$\mathscr{E}(g^h)=k(k-1)\sum_{p=0}^{k-1}(-1)^p\binom{k-1}{p}\int_0^1\left(\frac{w}{p+1}\right)^h(1-w)^{k-2}d\left(\frac{w}{p+1}\right)$$

と書ける．$w/(p+1)=y$ とおけば (17.6.28) は

(17.6.29)
$$\mathscr{E}(g^h)=k(k-1)\sum_{p=0}^{k-1}(-1)^p\binom{k-1}{p}\int_0^{\frac{1}{p+1}}y^h[1-(p+1)y]^{k-2}\,dy$$

となる．ここで

$$\theta_p(y)=\begin{cases}[1-(p+1)y]^{k-2}, & 0<y<\dfrac{1}{p+1}\\ 0, & \text{その他}\end{cases}$$

とすれば，$h=1,2,\cdots$ に対して

(17.6.30)
$$\mathscr{E}(g^h)=k(k-1)\int_0^1\sum_{p=0}^{k-1}(-1)^p\binom{k-1}{p}y^h\theta_p(y)\,dy$$

が成り立つ．

このようにして **5.5.1a** により g と y は同じ分布を持つことがわかる．したがって g の p.e. は

(17.6.31)　$$f(g)\,dg = k(k-1)\sum_{p=0}^{k-1}(-1)^p\binom{k-1}{p}\theta_p(g)\,dg.$$

$g>g'$ となる確率は $f(g)$ を積分して

(17.6.32)　$$P(g>g') = \sum_{p=0}^{r}(-1)^p\binom{k}{p+1}[1-(p+1)g']^{k-1}$$

となることが容易にわかる．ただし r は $1-(r+1)g'\geqq 0$ なる $k-1$ 以下の最大整数．この結果は最初に Fisher (1929) による異なった方法で得られた．k と α が与えられたとき，$P(g>g_\alpha)=\alpha$ なる g_α の値は有意水準 $100\alpha\%$ に対する g の棄却値である．Davis (1941) は $g_\alpha k=0.10(0.10)10.0$, $k=10(10)70$; $g_\alpha k=5.1(0.1)10.1$, k=80(10)160(20)300 について $P(g>g_\alpha)$ を作表している．

要約すると，次のようになる．

17.6.2　$x_1,\cdots,x_{2k+1}$ は独立な確率変数で分布 $N(0,\sigma^2)$ を持つとする．u_p, $p=1,\cdots,k$ を (17.6.19) で定義する．ただし $\alpha(\omega_p)$, $\alpha'(\omega_p)$ はそれぞれ (17.6.11), (17.6.12) で定義されている．このとき $g=\max\{u_p\}/(u_1+\cdots+u_k)$ ならば g の確率素分は (17.6.31) で与えられ，(0,1) 上の任意の g' に対する $P(g>g')$ は (17.6.32) で与えられる．

$u_1,\cdots,u_k$ を $u_1+\cdots+u_k$ で割って，その中で m 番目に大きい値を示す確率変数を g としたとき，上で述べた方法によって

(17.6.33)　$$P(g>g') = \frac{k!}{(m-1)!}\sum_{p=m}^{r}\frac{(-1)^{p-m}(1-pg')^{k-1}}{p(k-p)!(p-m)!}$$

が示される．ただし r は $1-rg'\geqq 0$ なる最大の整数である．

最後に，ここまでの議論においては x_t の平均を 0 と仮定していたことに注意しておこう．x_t の平均が未知の場合は，x_ξ を $x_\xi-\bar{x}$, $\xi=1,\cdots,2k+1$ に置き換えれば，少なくとも大きな n に対しては前述の結果に近い結果が得られる．

17.7　正規雑音の白色性検定

正規時系列 $\{x_t\}$ から標本 $(x_1,\cdots,x_n)$ を抽出したとする．このとき時系列が白色雑音

であるという仮説，すなわち，$x_1,\cdots,x_n$ は独立な確率変数で同一の正規分布 $N(\mu,\sigma^2)$ を持つという仮説を検定したい．この仮説検定について R. L. Anderson (1942) が提案した基準は，比

(17.7.1)　$$R=\frac{c_1'}{c_0'}$$

である．ただし

$$\sigma^2c_h' = \sum_{\xi=1}^{n}(x_\xi-\bar{x})(x_{\xi+h}-\bar{x}),\qquad h=0,1$$

で，$x_{n+1}\equiv x_1$ とする．$x_{n+1}\equiv x_1$ と置いて c_1' を輪環にしたのは t の 1 から $n-1$ までの和をとる場合に比較して勝手な操作と思われるかもしれないが，R に関する分布論を相当簡略化して取り扱うことができるからである．標本が白色雑音からならば n が大きいとき R は 0 に近づく傾向にある．そうでなければ R は 0 から離れた値を持つだろう．

さて c_1', c_0' は $x_1,\cdots,x_n$ の 2 次形式である．R' を定数とするとき差 $c_1'-R'c_0'$ も $x_1,\cdots,x_n$ の 2 次形式になっている．2 次形式 $c_1'-R'c_0'$ の分布を求めることによって R の分布を定めることができる．なぜなら

(17.7.2)　$$P(c_1'-R'c_0'>0)=P(R>R')$$

だからである．$(x_1,\cdots,x_n)$ が正規白色雑音ならば，$c_1'-R'c_0'$ の特性関数は

(17.7.3)　$$\varphi(t)=\left(\frac{1}{2\pi}\right)^{n/2}\int_{R_n}\exp\left[i(c_1'-R'c_0')t-\frac{1}{2\sigma^2}\sum_{\xi=1}^{n}(x_\xi-\mu)^2\right]dx_1\cdots dx_n$$

で与えられる．8.4 節の方法と結果を用い，$c_1'-R'c_0'$ が $(x_1-\mu),\cdots,(x_n-\mu)$ の 2 次形式であることに注意すれば

(17.7.4)　$$\varphi(t)=\prod_{\xi=1}^{n-1}[1-2it(d_\xi-R')]^{-\frac{1}{2}}$$

となる．ただし $d_\xi=\cos(2\pi\xi/n)$, $\xi=1,\cdots,n-1$．$d_\xi=d_{n-\xi}$ だから

(17.7.5)　$$\varphi(t)=\prod_{\xi=1}^{\frac{1}{2}(n-1)}[1-2it(d_\xi-R')]^{-1}$$

と書ける．ただし，ここでは簡単のために n を奇数としておく．$c_1'-R'c_0'=y$ の確率密度関数は **5.1.2** で与えられる．すなわち

(17.7.6)　$$g(y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-iyt}\varphi(t)\,dt.$$

さて，$\varphi(z)$ は複素平面において点 $z_\xi = -\frac{1}{2}i/(d_\xi - R'),\ \xi = 1, \cdots, \frac{n-1}{2}$ で 1 位の極を持つ．したがって (17.7.6) の積分を留数定理を用いて評価すると

$$g(y) = \frac{1}{2}\sum_{d_\xi > R'} A_\xi \exp\left\{-\frac{y}{2(d_\xi - R')}\right\} \tag{17.7.7}$$

が得られる．ただし

$$A_\xi = \frac{(d_\xi - R')^{\frac{1}{2}(n-5)}}{\prod_{j=1, j\neq\xi}^{\frac{1}{2}(n-1)} (d_\xi - d_j)} \tag{17.7.8}$$

よって $P(R > R') = P(y > 0)$ だから

$$P(R > R') = \int_0^\infty g(y)\,dy = \sum_{d_\xi > R'} A_\xi (d_\xi - R'). \tag{17.7.9}$$

ゆえに与えられた有意水準に対応する R' の棄却値が求められる．R. L. Anderson はこの方法を n が偶数の場合に拡張している．そこでは $\alpha = 0.99, 0.95, 0.05, 0.01,\ n = 5(1)15(5)75$ について，$P(R > R_\alpha) = \alpha$ なる R_α の値が作表されている．時系列に関する詳細な研究は T. W. Anderson (1948), Dixon (1944), Durbin と Watson (1951), Hannan (1955), Koopmans (1942), Moran (1948), Ogawara (1951), Quenouille (1948), Whittle (1951) 等により行なわれている．

17.8　時系列の線形予測

$x_t,\ t = \cdots, -1, 0, +1, \cdots$ はスペクトル密度関数 $f(\omega)$ を持つ時系列とする．このとき最小 2 乗法を用いてこの時系列の過去のすべての履歴から，x_1 を予測したい．これを数学的な問題として定式化すると次のようになる．いかなる条件の下で x_1 の最小 2 乗予測子が非決定的，すなわちこの予測子が正の 2 乗平均誤差を持つか，あるいは x_1 の最小 2 乗予測子が決定的，すなわち 0 の 2 乗平均誤差を持つのはどういう条件のときだろうか．この問題については Kolmogorov (1941) と Wiener (1949) が解を与えている．ここでは初等的な解析方法を用いて簡明に議論しよう．

履歴 $x_{-n}, \cdots, x_{-1}, x_0$ に基づく x_1 の予測子を $\hat{x}_{1,n}$ とすると，$\hat{x}_{1,n}$ は

$$\hat{x}_{1,n} = \sum_{\xi=-n}^{0} \hat{a}_{\xi,n} x_\xi \tag{17.8.1}$$

で与えられる．ただし $\hat{a}_{\xi,n},\ \xi = -n, \cdots, -1, 0$ は

$$\mathscr{E}\left(x_1 - \sum_{\xi=-n}^{0} a_{\xi,n} x_\xi\right)^2 \tag{17.8.2}$$

を最小にする $a_{\xi,n}$ の値である．3.8 節の最小 2 乗法を適用すれば

$$\hat{a}_{\xi,n} = \sum_{\xi'=-n}^{0} \sigma^{\xi\xi'} \gamma_{\xi'-1} \tag{17.8.3}$$

である．ただし $\|\sigma^{\xi\xi'}\| = \|\sigma_{\xi\xi'}\|^{-1}$ で $\|\sigma_{\xi\xi'}\|$ は $(n+1)\times(n+1)$ の行列

$$\begin{Vmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_n \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \gamma_n & \gamma_{n-1} & \cdots & \gamma_0 \end{Vmatrix}. \tag{17.8.4}$$

しかも γ_h は

$$\gamma_h = \gamma_{-h} = \int_{-\pi}^{+\pi} \cos h\omega\, f(\omega)\,d\omega \tag{17.8.5}$$

である．$\|\sigma_{\xi\xi'}\|$ の行列式を Δ_{n+1} で定義し，Δ_{n+2} も同様に定義すれば，(17.8.1) で与えられた予測子 $\hat{x}_{1,n}$ の最小 2 乗誤差は

$$\mathscr{E}(x_1 - \hat{x}_{1,n})^2 = \frac{\Delta_{n+2}}{\Delta_{n+1}} \tag{17.8.6}$$

である．さて $\xi = 1, \cdots, n$ に対して

$$\frac{\Delta_{n+2+\xi}}{\Delta_{n+1+\xi}} \leqslant \frac{\Delta_{n+2}}{\Delta_{n+1}} \leqslant \frac{\Delta_{n+2-\xi}}{\Delta_{n+1-\xi}}. \tag{17.8.7}$$

したがって

$$\prod_{\xi=1}^{n}\left(\frac{\Delta_{n+2+\xi}}{\Delta_{n+1+\xi}}\right) \leqslant \left(\frac{\Delta_{n+2}}{\Delta_{n+1}}\right)^n \leqslant \prod_{\xi=1}^{n}\left(\frac{\Delta_{n+2-\xi}}{\Delta_{n+1-\xi}}\right). \tag{17.8.8}$$

よって

$$\left(\frac{\Delta_{2n+2}}{\Delta_{n+2}}\right)^{1/n} \leqslant \frac{\Delta_{n+2}}{\Delta_{n+1}} \leqslant \left(\frac{\Delta_{n+1}}{\Delta_1}\right)^{1/n} \tag{17.8.9}$$

を得る．対数をとり $\frac{\Delta_{n+2}}{\Delta_{n+1}}$ を σ_n^2 で表わすと

$$A_n \leqslant \log \sigma_n^2 \leqslant B_n \tag{17.8.10}$$

となる．ただし

$$
(17.8.11)\qquad
\begin{aligned}
A_n &= \frac{1}{n}\sum_{\xi=1}^{2n+2}\log\theta_{\xi,2n+2} - \frac{1}{n}\sum_{\xi=1}^{n+2}\log\theta_{\xi,n+2}\\
B_n &= \frac{1}{n}\sum_{\xi=1}^{n+1}\log\theta_{\xi,n+1} - \frac{1}{n}\log\gamma_0
\end{aligned}
$$

で $(\theta_{1,2n+2},\cdots,\theta_{2n+2,2n+2})$ は行列式が Δ_{2n+2} の行列の特性根である．$(\theta_{1,n+2},\cdots,\theta_{n+2,n+2})$，$(\theta_{1,n+1},\cdots,\theta_{n+2,n+1})$ の定義も同様．この 3 つの行列はともに対角要素がすべて γ_0 に等しい正定値だから，各行列のすべての根は区間 $(0,\gamma_0)$ 上，したがってその対数は $(-\infty,\log\gamma_0)$ 上にある．

行列 (17.8.4) の特性根の対数の平均が，$n\longrightarrow\infty$ のとき，有限な極限値 K を持つならば

$$\lim_{n\to\infty}A_n=\lim_{n\to\infty}B_n=K$$

は明らかである．したがって

$$(17.8.12)\qquad \sigma^2=\lim_{n\to\infty}\mathscr{E}(x_1-\hat{x}_{1,n})^2$$

とおけば

$$(17.8.13)\qquad \sigma^2=e^K$$

となる．極限 K が有限でなければ，その値は $-\infty$ である．このとき $\sigma^2=0$．よって $\hat{x}_{1,n}$ は x_1 に対して $n\longrightarrow\infty$ のとき 0 に近づく平均 2 乗誤差を持つ．

さらに本書では議論しないが，より厳密な評価式を適用すれば，K が有限のときその値は

$$(17.8.14)\qquad K=\log 2\pi+\frac{1}{2\pi}\int_{-\pi}^{+\pi}\log f(\omega)\,d\omega$$

で与えられることが示されている．

この証明に興味がある読者は Kolmogorov (1941)，Wiener (1949) を参照されたい．Doob (1953)，Grenander と Rosenblatt (1957) にも証明されている．

時系列における線形予測に関する問題として他に**自己回帰方法**がある．これを簡単に述べよう．与えられた時系列 $\{x_t\}$ に対する r 次の自己回帰方法とは，定数 $a_1,\cdots,a_r$ に対して時系列

$$(17.8.15)\qquad y_t=x_t-a_1x_{t-1}-\cdots-a_rx_{t-r}$$

が白色雑音となることである．$a_1,\cdots,a_r$ の最小 2 乗推定量は標本 $(x_1,\cdots,x_n)$ からいつものように計算される．Dixon (1944) は y_t が白色雑音であることを仮定して尤度比による方法で，$a_m,a_{m+1},\cdots,a_r$ が 0 という対立仮説に対して，$a_1,\cdots,a_r$ が 0 という仮説を検定する問題を考えた．自己回帰に関する問題を別の立場から Mann と Wald (1943) は考察している．Wold (1938) は $\{y_t\}$ が白色雑音のとき時系列 $\{x_t\}$ が $y_t+a_1y_{t-1}+\cdots+a_ry_{t-r}$ なる型をしているかどうかを仮説検定している．

問　題

17.1 定常時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ はスペクトル密度関数

$$f(\omega)=\frac{1}{\pi^2}(\pi-|\omega|),\qquad -\pi<\omega<\pi$$

を持つとする．共分散関数 γ_h は

$$\gamma_h=\begin{cases}1, & h=0\\ \left(\dfrac{2}{\pi h}\right)^2, & h\ \text{が奇数}\\ 0, & h\ \text{が偶数}\end{cases}$$

であることを示せ．

17.2 (Slutzky (1937) の定理) x_t, $t=\cdots,-1,0,+1,\cdots$ はスペクトル密度関数 $f(\omega)$ を持つ実定常時系列とする．このとき平滑化された時系列

$$y_t=\sum_{p=0}^{k}a_px_{t-k+p},\qquad t=\cdots,-1,0,+1,\cdots$$

(ただし a は実定数) はスペクトル密度関数

$$\left[\sum_{q=0}^{k}\sum_{p=0}^{k}a_pa_q\cos(p-q)\omega\right]f(\omega)$$

を持つ定常時系列であることを示せ．特に $a_p=\dfrac{1}{k+1}$, $p=0,1,\cdots,k$ すなわち平滑化された過程が $k+1$ 個の成分の移動平均としてとられた場合，そのスペクトル密度関数は

$$\frac{1-\cos(k+1)\omega}{(k+1)^2(1-\cos\omega)}f(\omega)$$

となることを示せ．

17.3 (続き) 定常時系列 x_t, $t=\cdots,-1,0,+1,\cdots$ を前問のように r 回平滑化を繰り返せば，その時系列のスペクトル密度関数は

$$\left[\sum_{q=0}^{k}\sum_{p=0}^{k}a_pa_q\cos(p-q)\omega\right]^rf(\omega)$$

によることを示せ.

17.4　x_t,　$t = \cdots, -1, 0, +1, \cdots$ は分散 1 の白色雑音であるとする. この時系列の n 次（前向き）階差によって生成される時系列は共分散関数

$$\gamma_h = \begin{cases} (-1)^h \dfrac{(2n)!}{(n-h)!(n+h)!}, & -n \leqslant h \leqslant n \\ 0, & \text{その他} \end{cases}$$

を持つことを示せ.

17.5　x_t,　$t = \cdots, -1, 0, +1, \cdots$ は線形過程であるとする. すなわち時系列 x_t を

$$x_t = \sum_{p=-\infty}^{+\infty} a_{t-p} \varepsilon_p$$

とする. ただし各 a は $\sum_{p=-\infty}^{\infty} a_p^2 < +\infty$ を満たす実数で, 各 ε は平均 0, 分散 1 を持つ独立な確率変数である. この時系列のスペクトル密度関数 $f(\omega)$ は

$$f(\omega) = \frac{1}{2\pi} \sum_{q=-\infty}^{+\infty} \sum_{p=-\infty}^{+\infty} a_p a_{p+q} \cos q\omega$$

で与えられることを示せ.

17.6　x_t,　$t = \cdots, -1, 0, +1, \cdots$ は $\mathscr{E}(x_t) = 0$ なる定常時系列である. $F_M^*(\omega)$ を

$$F_M^*(\omega) = \frac{1}{\pi M} \int_{-\pi}^{\omega} \mathscr{E}(x_1 \sin\omega + \cdots + x_M \sin M\omega)^2 \, d\omega$$

で定義すれば, 時系列のスペクトル分布 $F(\omega)$ は $\lim_{M\to\infty} F_M^*(\omega)$ で与えられることを示せ. また $F_M^{**}(\omega)$ を

$$F_M^{**}\omega = \int_{-\pi}^{\omega} \mathscr{E} I_M(\omega) \, d\omega$$

とする. ただし $I_M(\omega)$ は

$$I_M(\omega) = \frac{1}{2\pi M} \left| \sum_{p=1}^{M} x_p e^{-ip\omega} \right|^2$$

で定義されるペリオドグラムである. $F(\omega)$ は $\lim_{M\to\infty} F_M^{**}(\omega)$ にも等しいことを示せ.

17.7　複素時系列 x_t,　$t = \cdots, -1, 0, +1, \cdots$ に対して共分散関数

$$\gamma_h = \mathscr{E}(x_{t+h}\bar{x}_t), \qquad h = \cdots, -1, 0, +1, \cdots$$

は (17.3.11) で与えられることを示せ. ただし $F(\omega) = \lim_{M\to\infty} F_M(\omega)$ で, $F_M(\omega)$ は (17.3.13) で定義されている.

17.8　x_t,　$t = \cdots, -1, 0, +1, \cdots$ が実時系列で, その共分散関数 γ_h は

$$\gamma_h = \gamma_0 \rho^h, \qquad 0 < \rho < +1, \qquad h = 0, 1, \cdots$$

であるとする. $x_{-n}, x_{-n+1}, \cdots, x_{-1}, x_0$ による x_1 の最小 2 乗線形推定量の分散を σ_n^2 とすれば, $\sigma_n^2 = \gamma_0(1-\rho^2)$ であることを示せ.

17.9（続き）　問題 17.8 における時系列のスペクトル密度関数は

$$f(\omega) = \frac{\gamma_0(1-\rho\cos\omega)}{\pi(1+\rho^2-2\rho\cos\omega)}$$

であることを示せ.

17.10　公式 (17.6.32) を導いたのと同様な方法によって, (17.6.33) を導け.

第18章　多変量解析

この 4 半世紀の間に急速に発展した数理統計学の分野の 1 つに多次元分布からの標本に対する 1 次および 2 次の解析がある．この分野は**多変量解析**としてよく知られるようになった．この分野の種々の問題に対して展開されてきた標本論や統計的仮説検定論の多くは多次元正規分布からの標本に関するものである．多変量解析に関する結果は，ほとんど数理統計学に関する文献に発表されてきた．本書のような性格の書物の中で，この分野における重要な興味ある結果をすべて網羅することは不可能である．したがって本章では多変量解析における本質的な概念と主要な結果だけに限定して議論せざるを得ない．さらに深く数学的な興味を持たれる読者は多変量解析に関する最近の著作 Anderson (1958) および Roy (1957) を参照されたい．Rao (1952)，Kendall (1957) には多変量解析の手法を数値問題に適用した詳しい例が見られる．Anderson (1958) には多変量解析に関する包括的な参考文献が記載されている．

18.1　多次元統計的散乱

多変量解析を紹介するにあたって，まず点のまわりの標本の**散乱**の定義をして議論しておくと便利であろう．

(a)　1次元の場合

まず $(x_1, \cdots, x_n)$ は 1 次元 c.d.f. $F(x)$ からの標本としよう．x_0 を任意の実数とする．$x_1, \cdots, x_n$ と x_0 の値が与えられれば，これらは実直線上の点として表わされる．このとき**ピボット点** x_0 のまわりの標本の散乱 ${}_1S_{x_0,n}$ は次の式で定義される．

(18.1.1)
$$ {}_1S_{x_0,n} = \sum_{\xi=1}^{n} (x_\xi - x_0)^2 = {}_1S_{\bar{x},n} + n(\bar{x} - x_0)^2. $$

${}_1S_{x_0,n}$ は x_0 と標本の各成分 x_ξ とによって生成された n 個の可能な $(x_\xi - x_0)$ の平方和にすぎない．x_0 のまわりの標本の散乱は平均 $\bar{x}$ のまわりの標本の散乱と x_0 のまわりの $\bar{x}$ の散乱を標本の大きさ n で加重した（n 倍した）量との和に等しい．このように平均のまわりの標本の散乱が最小になる．8.2 節より $F(x)$ が平均 μ，分散 σ^2 を持てば

(18.1.2)
$$ \mathscr{E}({}_1S_{x_0,n}) = n[\sigma^2 + (\mu - x_0)^2] $$

となる．よって $\mathscr{E}({}_1S_{x_0,n})$ の値は $x_0 = \mu$ を選んだときが最小になることは明らかである．

力学的用語で表現すれば ${}_1S_{x_0,n}$ はピボット点 x_0 のまわりの標本点 $x_1, \cdots, x_n$ の慣性能率ということになる．

さて 1 次元分布 $F_1(x)$，$F_2(x)$ からそれぞれ標本 $(x_1^{(1)}, \cdots, x_{n_1}^{(1)})$，$(x_1^{(2)}, \cdots, x_{n_2}^{(2)})$ を抽出する．$\bar{x}^{(1)}$，$\bar{x}^{(2)}$ をこの 2 つの標本の標本平均，$\bar{x}$ を合わせた標本の標本平均とする．${}_1S_{\bar{x}^{(1)},n_1}$，${}_1S_{\bar{x}^{(2)},n_2}$ はそれぞれ各平均のまわりの標本の散乱で，${}_1S_{\bar{x},n_1+n_2}$ はピボット点として $\bar{x}$ をとったときの $\bar{x}$ のまわりの合わせた標本の散乱である．ここで**合わせた標本**とは 2 つの標本をただ 1 つの（大きな）標本とみなしたときの大きさ $n_1 + n_2$ の標本を意味する．読者は

(18.1.3)
$$ {}_1S_{\bar{x},n_1+n_2} = {}_1S_{\bar{x}^{(1)},n_1} + {}_1S_{\bar{x}^{(2)},n_2} + n_1(\bar{x}^{(1)} - \bar{x})^2 + n_2(\bar{x}^{(2)} - \bar{x})^2 $$

を確かめることができるだろう．よって $\bar{x}$ のまわりの合わせた標本の散乱は各標本の成分が特定の値のとき最小値として各平均のまわりの 2 つの標本の散乱の和を持つ．この最小値は 2 つの標本平均が一致するときに限り起る．(18.1.3) で表わされた結果はただちに任意個数の標本の場合にも拡張できる．

(b)　2次元の場合

いま 2 次元 c.d.f. $F(x_1, x_2)$ から大きさ n の標本 $(x_{1\xi}, x_{2\xi},\ \xi = 1, \cdots, n)$ を考える．(x_{10}, x_{20}) は任意の実数の対とする．与えられた標本に対して標本は (x_{10}, x_{20}) とともに平面上の一群の点（高々 $n+1$ 個）として表現される．これを便宜上，**標本クラスター**と呼ぼう．標本の任意の 2 つの要素（対），たとえば $(x_{1\xi}, x_{2\xi})$ と $(x_{1\eta}, x_{2\eta})$ および点 (x_{10}, x_{20}) をとる．一般にこの 3 つの対は x_1x_2 平面内の 3 点として表わされる．3 点が同一直線上にないときは第 4 の点を，この点と与えられた 3 点とが平行 4 辺形の頂

点を構成するように 3 通りの方法で選ぶことができる．しかし 3 つのうち，どの方法によっても，構成された平行 4 辺形の面積はすべて $A_{\xi\eta,x_0}$ に等しい（ただし，面積の符号は考えない）．この面積の絶対値を $(x_{1\xi},x_{2\xi})$，$(x_{1\eta},x_{2\eta})$，(x_{10},x_{20}) によって決まる 2 次元容量と呼ぶ．この容量の平方は

(18.1.4)
$$A^2_{\xi\eta,x_0}=\begin{vmatrix} x_{1\xi}-x_{10} & x_{2\xi}-x_{20} \\ x_{1\eta}-x_{10} & x_{2\eta}-x_{20}\end{vmatrix}^2$$

である．右辺の行列式を平方すると，$A^2_{\xi\eta,x_0}$ は 4 つの行列式の和で表わされることがわかる．すなわち，

(18.1.5)
$$A^2_{\xi\eta,x_0}=\Delta_{\xi\xi}+\Delta_{\xi\eta}+\Delta_{\eta\xi}+\Delta_{\eta\eta}$$

ただし

(18.1.6)
$$\Delta_{\xi\eta}=\begin{vmatrix}(x_{1\xi}-x_{10})^2 & (x_{1\eta}-x_{10})(x_{2\eta}-x_{20}) \\ (x_{2\xi}-x_{20})(x_{1\xi}-x_{10}) & (x_{2\eta}-x_{20})^2\end{vmatrix}$$

で $\Delta_{\xi\xi}$，$\Delta_{\eta\xi}$，$\Delta_{\eta\eta}$ も同様に定義される．$\Delta_{\xi\xi}$ と $\Delta_{\eta\eta}$ はともに 0 である．

対 (ξ,η) は整数 $1,\cdots,n$ から選ばれた任意の 2 整数とすると $\binom{n}{2}$ 個の対がある．このような選び方全体にわたる $A^2_{\xi\eta,x_0}$ の和は $\sum_{\eta=1}^{n}\sum_{\xi=1}^{n}A^2_{\xi\eta,x_0}$ である．なぜなら，$A^2_{\xi\xi,x_0}=0$．しかし，この和は $\Delta_{\xi\eta}$（または $\Delta_{\eta\xi}$）の $\xi,\eta=1,\cdots,n$ にわたる和，すなわち $\sum_{\eta=1}^{n}\sum_{\xi=1}^{n}\Delta_{\xi\eta}$ に等しい．この和を ${}_2S_{x_0,n}$ で表わす．行列式の任意の集合 $\left\{\begin{vmatrix}a_\xi & b_\eta \\ c_\xi & d_\eta\end{vmatrix},\ \xi,\eta=1,\cdots,n\right\}$ に対して関係式

(18.1.7)
$$\sum_\eta\sum_\xi\begin{vmatrix}a_\xi & b_\eta \\ c_\xi & d_\eta\end{vmatrix}=\begin{vmatrix}\sum_\xi a_\xi & \sum_\eta b_\eta \\ \sum_\xi c_\xi & \sum_\eta d_\eta\end{vmatrix}$$

が成り立つことに注意すれば，結局

(18.1.8)
$${}_2S_{x_0,n}=\begin{vmatrix}\sum_\xi(x_{1\xi}-x_{10})^2 & \sum_\eta(x_{1\eta}-x_{10})(x_{2\eta}-x_{20}) \\ \sum_\xi(x_{2\xi}-x_{20})(x_{1\xi}-x_{10}) & \sum_\eta(x_{2\eta}-x_{20})^2\end{vmatrix}$$

が成立する．ただし $\sum_\xi$，$\sum_\eta$ はともに $1,\cdots,n$ にわたる和である．

行列式が ${}_2S_{x_0,n}$ である行列はグラミアン行列であるともいう．すなわち行列 A が

(18.1.9)
$$\begin{Vmatrix}x_{11}-x_{10} & x_{21}-x_{20} \\ \vdots & \vdots \\ x_{1n}-x_{10} & x_{2n}-x_{20}\end{Vmatrix}$$

で，A の転置行列を A' とするとき $A'A$ なる形の行列をグラミアン行列という．

量 ${}_2S_{x_0,n}$ をピボット点 (x_{10},x_{20}) のまわりの標本 $(x_{1\xi},x_{2\xi},\ \xi=1,\cdots,n)$ の散乱という．これは (x_{10},x_{20}) と標本 $(x_{1\xi},x_{2\xi},\ \xi=1,\cdots,n)$ のすべての可能な対とで決まる 2 次元容量の平方和である，行列式が ${}_2S_{x_0,n}$ となるような行列は (x_{10},x_{20}) のまわりの標本の散乱行列と呼ばれる．

さて $\bar{x}_i=\dfrac{1}{n}\sum_{\xi=1}^{n}x_{i\xi}$，$i=1,2$ で

(18.1.10)
$$u_{ij}=u_{ji}=\sum_{\xi=1}^{n}(x_{i\xi}-\bar{x}_i)(x_{j\xi}-\bar{x}_j),\qquad i,j=1,2$$

とすれば

(18.1.11)
$$\sum_{\xi=1}^{n}(x_{i\xi}-x_{i0})(x_{j\xi}-x_{j0})=u_{ij}+n(\bar{x}_i-x_{i0})(\bar{x}_j-x_{j0})$$

である．したがって

(18.1.12)
$${}_2S_{x_0,n}=\begin{vmatrix}u_{11}+n(\bar{x}_1-x_{10})^2 & u_{12}+n(\bar{x}_1-x_{10})(\bar{x}_2-x_{20}) \\ u_{21}+n(\bar{x}_2-x_{20})(\bar{x}_1-x_{10}) & u_{22}+n(\bar{x}_2-x_{20})^2\end{vmatrix}.$$

これは (18.1.1) の 2 次元の場合に相当する．右辺の行列式を 4 つの行列の和で書き表わせば ${}_2S_{x_0,n}$ は

(18.1.13)
$${}_2S_{x_0,n}=\begin{vmatrix}u_{11} & u_{12} \\ u_{21} & u_{22}\end{vmatrix}\cdot\left[1+n\sum_{i,j=1}^{2}u^{ij}(\bar{x}_i-x_{i0})(\bar{x}_j-x_{j0})\right]$$

になる．ただし $\|u^{ij}\|=\|u_{ij}\|^{-1}$．もちろんこの行列は $(\bar{x}_1,\bar{x}_2)$ のまわりの散乱行列 $\|u_{ij}\|$ が確率 1 で正則であるときに限り存在する．便宜上，$\|u_{ij}\|$ を内部散乱行列，この行列式 $|u_{ij}|$ を標本の内部散乱と呼ぼう．$\|u_{ij}\|$ が正則であるための必要十分条件は標本の n 個の点 $(x_{1\xi},x_{2\xi},\ \xi=1,\cdots,n)$ が同一直線上にないことである．このように $\|u_{ij}\|$ が正則ならば (18.1.13) における右辺の［　］内の 2 次形式は正定値である．したがって，この場合は明らかに次が成り立つ．

18.1.1　(x_{10},x_{20}) のまわりの標本の散乱を最小にするピボット点 (x_{10},x_{20}) は標本平均 $(\bar{x}_1,\bar{x}_2)$ である．このとき最小散乱 ${}_2S_{\bar{x},n}$ は

(18.1.14)
$${}_2S_{\bar{x},n}=\begin{vmatrix}u_{11} & u_{12} \\ u_{21} & u_{22}\end{vmatrix}$$

で与えられる内部散乱である．

次も示される.

18.1.2　(x_{10}, x_{20}) のまわりの 2 次元散乱の平均値は

(18.1.15)

$$\mathscr{E}({}_2S_{x_0,n}) = 2\binom{n}{2}\begin{vmatrix} \sigma_{11}+(\mu_1-x_{10})^2 & \sigma_{12}+(\mu_1-x_{10})(\mu_2-x_{20}) \\ \sigma_{21}+(\mu_2-x_{20})(\mu_1-x_{10}) & \sigma_{22}+(\mu_2-x_{20})^2 \end{vmatrix}$$

で与えられる. ただし $\|\sigma_{ij}\|$, (μ_1, μ_2) は $F(x_1, x_2)$ に関するそれぞれ共分散行列, 平均ベクトル. $\mathscr{E}({}_2S_{x_0,n})$ が最小になるのは $(x_{10}, x_{20}) = (\mu_1, \mu_2)$ のときであり, このときに限る. この最小値は $\min\limits_{(x_{10},x_{20})} \mathscr{E}({}_2S_{x_0,n}) = \mathscr{E}({}_2S_{\mu,n}) = 2\binom{n}{2}|\sigma_{ij}|$ である.

公式 (18.1.15) は (18.1.2) の 2 次元の場合に相当する. この結果を示すには, まず $\mathscr{E}(\Delta_{\xi\eta})$ を

(18.1.16)
$$\mathscr{E}(\Delta_{\xi\eta}) = \int_{R_2}\int_{R_2} \Delta_{\xi\eta}\, dF(x_{1\xi}, x_{2\xi})\, dF(x_{1\eta}, x_{2\eta})$$

で定義しておこう. (18.1.7) は有限和の場合はもちろん積分に対しても成り立つから (18.1.16) の右辺を行列式 $|D_{ij}|$ で表わすことができる. ただし

$$D_{21} = \int_{R_2} (x_{2\xi} - x_{20})(x_{1\xi} - x_{10})\, dF(x_{1\xi}, x_{2\xi})$$

で, 他の D_{11}, D_{12}, D_{22} についても同様である. すなわち

$$D_{ij} = \mathscr{E}(x_i - x_{i0})(x_j - x_{j0}) = \sigma_{ij} + (\mu_i - x_{i0})(\mu_j - x_{j0}).$$

よって $\xi \neq \eta = 1, 2, \cdots, n$ のとき

$$\mathscr{E}(\Delta_{\xi\eta}) = \begin{vmatrix} \sigma_{11}+(\mu_1-x_{10})^2 & \sigma_{12}+(\mu_1-x_{10})(\mu_2-x_{20}) \\ \sigma_{21}+(\mu_2-x_{20})(\mu_1-x_{10}) & \sigma_{22}+(\mu_2-x_{20})^2 \end{vmatrix}$$

しかし $\mathscr{E}(A^2_{\xi\eta,x_0}) = \mathscr{E}(\Delta_{\xi\xi}) + \mathscr{E}(\Delta_{\xi\eta}) + \mathscr{E}(\Delta_{\eta\xi}) + \mathscr{E}(\Delta_{\eta\eta}) = 2\mathscr{E}(\Delta_{\xi\eta})$. さらに ${}_2S_{x_0,n}$ は $\binom{n}{2}$ 個の $A^2_{\xi\eta,x_0}$ の和からなっているから

$$\mathscr{E}({}_2S_{x_0,n}) = 2\binom{n}{2}\mathscr{E}(\Delta_{\xi\eta}).$$

これは (18.1.15) の右辺に等しい. これで **18.1.2** の証明を終える.

$\|\sigma_{ij}\|$ が正定値ならば $\|\sigma^{ij}\| = \|\sigma_{ij}\|^{-1}$ が存在して, これも正定値である. ゆえに

(18.1.17)
$$\mathscr{E}({}_2S_{x_0,n}) = 2\binom{n}{2}|\sigma_{ij}|\left[1 + \sum_{i,j=1}^{2} \sigma^{ij}(\mu_i - x_{i0})(\mu_j - x_{j0})\right]$$

となり, (18.1.17) から次が明らかになる.

18.1.3　$\|\sigma_{ij}\|$ が正定値ならば, $\mathscr{E}({}_2S_{x_0,n})$ は (x_{10}, x_{20}) として $F(x_1, x_2)$ に関する平均値 (μ_1, μ_2) を選んだとき最小になる. この最小値は

(18.1.18)
$$\mathscr{E}({}_2S_{\mu,n}) = \mathscr{E}|v_{ij}| = 2\binom{n}{2}|\sigma_{ij}|.$$

ただし

(18.1.19)
$$v_{ij} = v_{ji} = \sum_{\xi=1}^{n} (x_{i\xi} - \mu_i)(x_{j\xi} - \mu_j), \qquad i, j = 1, 2.$$

いま, $(x^{(1)}_{1\xi_1}, x^{(1)}_{2\xi_1},\ \xi_1 = 1, \cdots, n_1)$, $(x^{(2)}_{1\xi_2}, x^{(2)}_{2\xi_2},\ \xi_2 = 1, \cdots, n_2)$ をそれぞれ c.d.f. $F_1(x_1, x_2)$, $F_2(x_1, x_2)$ からの標本とする. $(\bar{x}^{(1)}_1, \bar{x}^{(1)}_2)$, $(\bar{x}^{(2)}_1, \bar{x}^{(2)}_2)$ は各標本の平均で $\|u^{(1)}_{ij}\|$, $\|u^{(2)}_{ij}\|$, $i, j = 1, 2$ はそれぞれ標本の内部散乱行列である. 2つの標本を合わせて, その標本平均を $(\bar{x}_1, \bar{x}_2)$ とすれば $(\bar{x}_1, \bar{x}_2)$ のまわりの合わせた標本の散乱は

(18.1.20)
$$\begin{aligned} &{}_2S_{\bar{x},n_1+n_2} \\ &\quad = |u^{(1)}_{ij} + u^{(2)}_{ij} + n_1(\bar{x}^{(1)}_i - \bar{x}_i)(\bar{x}^{(1)}_j - \bar{x}_j) + n_2(\bar{x}^{(2)}_i - \bar{x}_i)(\bar{x}^{(2)}_j - \bar{x}_j)| \end{aligned}$$

で与えられる. (18.1.20) における右辺の行列式は (18.1.3) の右辺の 2 次元の場合に相当する.

(c) k 次元の場合

k 次元 c.d.f. $F(x_1, \cdots, x_k)$ からの標本 $(x_{1\xi}, \cdots, x_{k\xi}, \xi = 1, \cdots, n)$ のピボット点 $(x_{10}, \cdots, x_{k0})$ のまわりの散乱 ${}_kS_{x_0,n}$ を定義するのは難しくない. 標本とピボット点はユークリッド空間 R_k における $n+1$ 個の点として表わされる. これを**標本クラスター**と呼ぶ. k 個の異なった標本点, たとえば $(x_{1\xi_i}, \cdots, x_{k\xi_i})$, $i = 1, \cdots, k$ とピボット点 $(x_{10}, \cdots, x_{k0})$ に対してはすでに述べたように R_k の点を選んで, この点と $k+1$ 個の点によって k 次元平行体を構成するようにできる. このような点の選び方は $k+1$ 通りある. しかし $k+1$ 個の異なった方法により生成される平行体の k 次元体積はいずれも等しい (ただし符号は考えない). この絶対値を $(x_{1\xi_i}, \cdots, x_{k\xi_i})$, $i = 1, \cdots, k$ と $(x_{10}, \cdots, x_{k0})$ に

よって定まる k 次元容量という．したがって散乱 ${}_kS_{x_0,n}$ は $(x_{10},\cdots,x_{k0})$ と $\binom{n}{k}$ 通りの異なる $(x_{1\xi_i},\cdots,x_{k\xi_i})$, $i=1,\cdots,k$ の選び方のおのおのによって定まる k 次元容量の2乗和として定義される．${}_kS_{x_0,n}$ は明らかに (18.1.11) で与えられた ${}_2S_{x_0,n}$ の k 次元への拡張である．すなわち

$$(18.1.21)\qquad {}_kS_{x_0,n}=|u_{ij}+n(\bar{x}_i-x_{i0})(\bar{x}_j-x_{j0})|.$$

ここで u_{ij}, $i,j=1,\cdots,k$ の定義は (18.1.10) から明らかである．標本の内部散乱行列 $\|u_{ij}\|$ が正定値となるための必要十分条件は n 個の標本点 $(x_{1\xi},\cdots,x_{k\xi},\ \xi=1,\cdots,n)$ が k 次元より低次元の超平面上にないことである．$\|u_{ij}\|$ が正定値ならば ${}_kS_{x_0,n}$ は次のように書き改められる．

$$(18.1.22)\qquad {}_kS_{x_0,n}=|u_{ij}|\cdot\left[1+n\sum_{i,j=1}^{k}u^{ij}(\bar{x}_i-x_{i0})(\bar{x}_j-x_{j0})\right]$$

ただし $\|u^{ij}\|=\|u_{ij}\|^{-1}$ で $\|u_{ij}\|$ が正定値ならば $\|u^{ij}\|$ も正定値．(18.1.22) は (18.1.13) の k 次元への拡張になっている．[ ] 内の2次形式は正定値だから k 次元の場合については **18.1.1** に相当する定理が得られる．これは次のようになる．$(x_{10},\cdots,x_{k0})$ として標本平均ベクトル $(\bar{x}_1,\cdots,\bar{x}_k)$ を選んだとき ${}_kS_{x_0,n}$ は最小になる．このとき ${}_kS_{x_0,n}$ の最小値は ${}_kS_{\bar{x},n}$ である．ただし

$$(18.1.23)\qquad {}_kS_{\bar{x},n}=|u_{ij}|,\qquad i,j=1,\cdots,k.$$

もちろんこれは標本の内部散乱である．

${}_kS_{x_0,n}$ は A を $n\times k$ 行列

$$\begin{Vmatrix}(x_{11}-x_{10})\cdots & (x_{k1}-x_{k0})\\ \vdots & \vdots\\ (x_{1n}-x_{10})\cdots & (x_{kn}-x_{k0})\end{Vmatrix}$$

としたときグラミアン行列 $A'A$ の行列式になっている．ただし A' は A の転置行列．また $\|u_{ij}\|$ もグラミアン行列 $B'B$ になる．ただし B は A において x_{i0} を $\bar{x}_i$, $i=1,\cdots,k$ で置き換えた行列で B' はその転置行列．

k 次元分布からの標本 $(x_{1\xi},\cdots,x_{k\xi},\ \xi=1,\cdots,n)$ の平均値ベクトルを $(\mu_1,\cdots,\mu_k)$，(正則な) 分散行列を $\|\sigma_{ij}\|$ とすれば (18.1.15) を導いた方法によって

$$(18.1.24)\qquad \mathscr{E}({}_kS_{x_0,n})=k!\binom{n}{k}\cdot|\sigma_{ij}+(\mu_i-x_{i0})(\mu_j-x_{j0})|$$

が示される．さらに次のように書き直される．

$$(18.1.25)\qquad \mathscr{E}({}_kS_{x_0,n})=k!\binom{n}{k}\cdot|\sigma_{ij}|\cdot\left[1+\sum_{i,j=1}^{k}\sigma^{ij}(\mu_i-x_{i0})(\mu_j-x_{j0})\right].$$

これは (18.1.17) の k 次元の場合に相当する．よって **18.1.2** は k 次元の場合には次のようになる．$\mathscr{E}({}_kS_{x_0,n})$ が最小となるのは $(x_{10},\cdots,x_{k0})$ として平均値ベクトル $(\mu_1,\cdots,\mu_k)$ を選んだときであり，かつ，このときに限る．この最小値は $\mathscr{E}({}_kS_{\mu,n})$ で次の値を持つ．

$$(18.1.26)\qquad \mathscr{E}({}_kS_{\mu,n})=\mathscr{E}|v_{ij}|=k!\binom{n}{k}|\sigma_{ij}|$$

ただし

$$(18.1.27)\qquad v_{ij}=v_{ji}=\sum_{\xi=1}^{n}(x_{i\xi}-\mu_i)(x_{j\xi}-\mu_j).$$

行列式 $|\sigma_{ij}|$ は $F(x_1,\cdots,x_k)$ の**一般化分散**と呼ばれている．これはまた c.d.f. $F(x_1,\cdots,x_k)$ を持つ分布の内部散乱ともみなされる．

$k=2$ のとき (18.1.20) が完全に示されれば (18.1.20) の次の k 次元への拡張も容易に導かれるであろう．合わせた標本の標本平均値のまわりの散乱 ${}_kS_{\bar{x},n_1+n_2}$ は (18.1.20) (ただし $i,j=1,\cdots,k$) で与えられる．多次元統計的散乱に関する詳細な議論は Wilks (1960 *a*) を参照されたい．

18.2　ウィシャート分布

(a)　ウィシャート分布の導入

前節では，有限な平均値と共分散行列を持つ任意の分布からの標本に対しては，k 次元分布からの標本散乱の平均値が比較的容易に求められることを見てきた．しかし k 次元正規分布からの標本の場合をのぞけば，散乱の高次のモーメントを求める問題は，このような一般的分布からの標本の場合には大変難しくなる．事実，k 次元正規分布の平均値をピボット点 $(x_{10},\cdots,x_{k0})$ に選んだときには，この正規分布からの標本に対する散乱のモーメントばかりでなく散乱行列の要素の分布関数も求めることができる．この散乱行列の要素の分布は **Wishart** (1928) 分布としてよく知られている．この分布は，これから展開される多変量統計的推定論において基本的な役割を果たしている．基本的には特性関数

を用いてウィシャート分布を導こう．この特性関数は $\|\sigma_{ij}\|$ が正定値で $k \leqq n$ ならば，$\|v_{ij}\|$ が正定値となるような $\{v_{ij}\}$ の空間のすべての点で意味を持つことがわかるであろう．

$(x_{1\xi}, \cdots, x_{k\xi},\ \xi = 1, \cdots, n)$ は p.d.f. (7.4.1) を持つ k 次元正規分布 $N(\{\mu_i\}\,;\,\|\sigma_{ij}\|)$ からの標本とする．k 次元分布からの大きさ n の標本は 8.1 節で定義されている．この標本の母集団 $(\mu_1, \cdots, \mu_k)$ のまわりの散乱 ${}_kS_{\mu,n}$ を考える．

(18.2.1) $$ {}_kS_{\mu,n} = |v_{ij}| $$

である．ただし

(18.2.2) $$ v_{ij} = v_{ji} = \sum_{\xi=1}^{n} (x_{i\xi} - \mu_i)(x_{j\xi} - \mu_j), \qquad i, j = 1, \cdots, k. $$

$\{v_{ii},\ i = 1, \cdots, k\,;\, 2v_{ij},\ i > j = 1, \cdots, k\}$ の特性関数を $\varphi(\{t_{ij}\})$，$t_{ij} \equiv t_{ji}$ で表わせば

(18.2.3) $$ \varphi(\{t_{ij}\}) = \left[\frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{k/2}}\right]^n \int_{R_{nk}} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{k} \sigma^{ij} v_{ij} + \mathrm{i}\sum_{i,j=1}^{k} v_{ij} t_{ij}\right] \prod_{i,\xi} dx_{i\xi} $$

となる．$\varphi(\{0\})$*) $= 1$ より (18.2.3) の積分値は $[(2\pi)^{\frac{1}{2}k}/\sqrt{|\sigma^{ij} - 2\mathrm{i}t_{ij}|}]^n$ を持つ．よって

(18.2.4) $$ \varphi(\{t_{ij}\}) = |\tau_{ij}|^{\frac{1}{2}n} \cdot |\tau_{ij} - \mathrm{i}t_{ij}|^{-\frac{1}{2}n}. $$

ただし

(18.2.5) $$ \tau_{ij} = \frac{1}{2}\sigma^{ij}, \qquad i, j = 1, \cdots k. $$

多次元のレヴィの定理，すなわち **5.2.1** を適用すれば $\{v_{ii},\ i = 1, \cdots, k\,;\, 2v_{ij},\ i > j = 1, \cdots, k\}$ の確率密度関数は

(18.2.6) $$ f(\{v_{ii}, 2v_{ij}\}) = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}k(k+1)} \int_{R_{\frac{1}{2}k(k+1)}} \exp\left[-\mathrm{i}\sum_{i,j=1}^{k} v_{ij} t_{ij}\right] \varphi(t_{ij}) \prod_{i \geqq j=1}^{k} dt_{ij} $$

で与えられる．

いま，$\tau_{ij} - \mathrm{i}t_{ij} = \theta_{ij}$ とおくと $\theta_{ij} = \theta_{ji}$，$-\mathrm{i}t_{ij} = \theta_{ij} - \tau_{ij}$，$dt_{ij} = \mathrm{i}d\theta_{ij}$ だから

(18.2.7) $$ f(\{v_{ii}, 2v_{ij}\}) = A \int_{\substack{\tau_{ij}-\mathrm{i}\infty \\ i \geqq j=1,\cdots,k}}^{\tau_{ij}+\mathrm{i}\infty} \exp\left[\sum_{i,j=1}^{k} v_{ij}\theta_{ij}\right] \cdot |\theta_{ij}|^{-\frac{1}{2}n} \prod_{i \geqq j=1}^{k} d\theta_{ij} $$

を得る．ただし

*) 0 は 0 ベクトルを意味する．［訳注］

(18.2.8) $$ A = \frac{|\tau_{ij}|^{\frac{1}{2}n} \exp\left[-\sum_{i,j=1}^{k} \tau_{ij} v_{ij}\right]}{(2\pi\mathrm{i})^{\frac{1}{2}k(k+1)}}. $$

次のように逐次，積分を行なおう．まず θ_{kk} で積分してから $\theta_{1k}, \cdots, \theta_{k-1,k}$ で積分する．次に，$\theta_{k-1,k-1}$ に関して行ない，さらに $\theta_{1,k-1}, \cdots, \theta_{k-2,k-1}$ で行なう．この方法を k 回繰り返す．このとき (18.2.7) は次のようになる．

(18.2.9) $$ f(\{v_{ii}, 2v_{ij}\}) = AH_1 \int_{\substack{\tau_{ik}-\mathrm{i}\infty \\ i=1,\cdots,k-1}}^{\tau_{ik}+\mathrm{i}\infty} \exp\left[2\sum_{i=1}^{k-1} v_{ik}\theta_{ik}\right] H_2 \prod_{i=1}^{k-1} d\theta_{ik} $$

ここに

(18.2.10)
$$ H_1 = \int_{\tau_{ij}-\mathrm{i}\infty}^{\tau_{ij}+\mathrm{i}\infty} \exp\left[\sum_{i,j=1}^{k-1} v_{ij}\theta_{ij}\right] \cdot |\theta_{ij}|_{(k-1)}^{-\frac{1}{2}n} \prod_{i \geqq j=1}^{k-1} d\theta_{ij} $$
$$ H_2 = \int_{\tau_{kk}-\mathrm{i}\infty}^{\tau_{kk}+\mathrm{i}\infty} \exp(v_{kk}\theta_{kk}) \cdot \left[\theta_{kk} - \sum_{i,j=1}^{k-1} \theta^{ij}_{(k-1)}\theta_{ik}\theta_{jk}\right]^{-\frac{1}{2}n} d\theta_{kk} $$

で

$$ \|\theta^{ij}_{(k-1)}\| = \|\theta_{ij}\|^{-1}_{(k-1)}, \qquad \|\theta_{ij}\|_{(k-1)} = \begin{Vmatrix} \theta_{11} & \cdots & \theta_{1,k-1} \\ \vdots & & \vdots \\ \theta_{k-1,1} & \cdots & \theta_{k-1,k-1} \end{Vmatrix} $$

である．さて H_2 において

(18.2.11) $$ \theta_{kk} - \sum_{i,j=1}^{k-1} \theta^{ij}_{(k-1)}\theta_{ik}\theta_{jk} = w $$

とすれば

(18.2.12) $$ H_2 = \exp\left[v_{kk}\sum_{i,j=1}^{k-1} \theta^{ij}_{(k-1)}\theta_{ik}\theta_{jk}\right] \cdot \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} e^{v_{kk}w}\, w^{-\frac{1}{2}n}\, dw $$

が成り立つ．ただし c は実数で正の定数であるが，ここではその値を書くにはおよばない．(18.2.12) における積分は基本的にはハンケル積分 (5.1 節の例を参照) で，その値 I_1 は

(18.2.13) $$ I_1 = \frac{(v_{kk})^{\frac{1}{2}(n-1)}(2\pi\mathrm{i})}{\Gamma\left(\frac{1}{2}n\right)} $$

で与えられる．H_2 の値を (18.2.9) に代入すると

(18.2.14) $$ f(\{v_{ii}, 2v_{ij}\}) = A \cdot H_1 \cdot I_1 \cdot G $$

を得る．ただし

(18.2.15)　$$G = \int_{\substack{\tau_{ik}-i\infty \\ i=1,\cdots,k-1}}^{\tau_{ik}+i\infty} \exp\left[v_{kk}\sum_{i,j=1}^{k-1}\theta^{ij}_{(k-1)}\theta_{ik}\theta_{jk} + \sum_{i=1}^{k-1} v_{ik}\theta_{ik}\right]\prod_{i=1}^{k-1} d\theta_{ik}.$$

G に変換 $\theta_{ik} = i\phi_{ik}$, $i = 1, \cdots, k-1$ を施せば

(18.2.16)　$$G = (i)^{k-1}\int_{\substack{-\infty+i\tau_{ik} \\ i=1,\cdots,k-1}}^{+\infty+i\tau_{ik}} \exp\left[-v_{kk}\sum_{i,j=1}^{k-1}\theta^{ij}_{(k-1)}\phi_{ik}\phi_{jk} + 2i\sum_{i=1}^{k-1} v_{ik}\phi_{ik}\right]\prod_{i=1}^{k-1} d\phi_{ik}$$

になる． (18.2.16) における積分は $(k-1)$ 複素変数関数の $(k-1)$ 次元複素空間における虚数部の値をすべて固定した実部の全 $(k-1)$ 次元空間にわたる積分である．この積分は虚数部の固定された値の選び方によらないことが示される．したがって虚数部をすべて 0 に固定してもよい．よって積分は $\phi_{1k}, \cdots, \phi_{k-1,k}$ を実数としてもよく，$\phi_{1k}, \cdots, \phi_{k-1,k}$ の $(k-1)$ 次元空間にわたる積分に等しい． この積分値は (7.4.13), (7.4.16) を考慮すれば次の値になる．

(18.2.17)　$$G = J_1 \cdot |\theta^{ij}_{(k-1)}|^{-\frac{1}{2}} \exp\left[-\sum_{i,j=1}^{k-1}\theta_{ij}\frac{v_{ik}v_{jk}}{v_{kk}}\right]$$

ただし

(18.2.18)　$$J_1 = \frac{\pi^{\frac{1}{2}(k-1)} i^{k-1}}{(v_{kk})^{\frac{1}{2}(k-1)}}.$$

G の値を (18.2.14) に代入すれば次のように書き表わせる．

(18.2.19)　$$f(\{v_{ii}, 2v_{ij}\}) = A(I_1J_1)\int_{\substack{\tau_{ij}-i\infty \\ i\geqq j=1,\cdots,k-1}}^{\tau_{ij}+i\infty} \exp\left[\sum_{i,j=1}^{k-1} v^{(1)}_{ij}\theta_{ij}\right] \cdot |\theta_{ij}|^{\frac{1}{2}(n-1)}_{(k-1)} \prod_{i\geqq j=1}^{k-1} d\theta_{ij}.$$

ただし

(18.2.20)　$$v^{(1)}_{ij} = v_{ij} - \frac{v_{ik}v_{jk}}{v_{kk}}, \qquad i, j = 1, \cdots, k-1.$$

(18.2.19) における積分は k, n, v_{ij} をそれぞれ $k-1$, $n-1$, $v^{(1)}_{ij}$ に置き換えれば (18.2.7) の積分と同じ型をしている．

したがって，これまで述べたような積分を逐次繰り返せば

(18.2.21)　$$f(\{v_{ii}, 2v_{ij}\}) = A(I_1J_1)(I_2J_2)\cdots(I_kJ_k)$$

を得る．ただし I_1, J_1 は (18.2.13), (18.2.18) で与えられるが，他の I, J は

(18.2.22)
$$I_2 = \frac{(v^{(1)}_{k-1,k-1})^{\frac{1}{2}(n-1)-1}(2\pi i)}{\Gamma\left(\frac{n-1}{2}\right)}, \cdots, I_k = \frac{(v^{(k-1)}_{11})^{\frac{1}{2}(n-k+1)-1}(2\pi i)}{\Gamma\left(\frac{n-k+1}{2}\right)}$$
$$J_2 = \frac{\pi^{\frac{1}{2}(k-2)} i^{k-2}}{(v^{(1)}_{k-1,k-1})^{\frac{1}{2}(k-2)}}, \cdots, J_k = \frac{\pi^{\frac{1}{2}(k-k)} i^{k-k}}{(v^{(k-1)}_{11})^{\frac{1}{2}(k-k)}} = 1$$

である．$v^{(1)}_{ij}$ は (18.2.20) で与えられ

(18.2.23)
$$v^{(2)}_{ij} = v^{(1)}_{ij} - \frac{v^{(1)}_{i,k-1}v^{(1)}_{j,k-1}}{v^{(1)}_{k-1,k-1}}, \qquad i, j = 1, \cdots, k-2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$v^{(k)}_{11} = v^{(k-1)}_{11} - \frac{v^{(k-1)}_{12}v^{(k-1)}_{12}}{v^{(k-1)}_{22}}$$

である．$I_1, \cdots, I_k$, $J_1, \cdots, J_k$ の値を (18.2.21) に代入すれば，

(18.2.24)　$$f(\{v_{ii}, 2v_{ij}\}) = \frac{|\tau_{ij}|^{\frac{1}{2}n}[v_{kk}v^{(1)}_{k-1,k-1}\cdots v^{(k-1)}_{11}]^{\frac{1}{2}(n-k-1)} \exp\left[-\sum_{i,j=1}^{k}\tau_{ij}v_{ij}\right]}{2^{\frac{1}{2}k(k-1)}\pi^{k(k-1)/4}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-1}{2}\right)\cdots\Gamma\left(\frac{n-k+1}{2}\right)}$$

は $\|v_{ij}\|$ が正定値となる $\{v_{ii}, 2v_{ij}\}$ の標本空間の任意の点に対して成り立つ． その他の v の点では $f(\{v_{ii}, 2v_{ij}\}) = 0$ である． しかし列 $v_{kk}, v^{(1)}_{k-1,k-1}, \cdots, v^{(k-1)}_{11}$ の要素は 7.4 (a) 節で定義された列 $\sigma^{11}, \sigma^{22}_{(1)}, \cdots, \sigma^{kk}_{(k-1)}$ の要素と同じ構造をしている． よって (7.4.6) から

(18.2.25)　$$v_{kk}v^{(1)}_{k-1,k-1}\cdots v^{(k-1)}_{11} = |v_{ij}|$$

を得る．(18.2.24) の右辺の表現は $\{v_{ii},\ i = 1, \cdots, k\,;\ 2v_{ij},\ i > j = 1, \cdots, k\}$ の p.d.f. である．したがって $g(\{v_{ij}\})$ を $\{v_{ij},\ i \geqq j = 1, \cdots, k\}$ の p.d.f. とすれば

(18.2.26)　$$g(\{v_{ij}\}) = 2^{\frac{1}{2}k(k-1)}f(\{v_{ii}, 2v_{ij}\})$$

が得られる．

$\tau_{ij} = \dfrac{1}{2}\sigma^{ij}$ を考慮すれば結局 Wishart (1928) による次の基本的な結果を得る．

18.2.1　$(x_{1\xi}, \cdots, x_{k\xi},\ \xi = 1, \cdots, n)$, $k \leqq n$ が k 次元分布 $N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ からの標本で，$\|v_{ij}\|$ を (18.2.2) で定義された母（集団）平均 $(\mu_1, \cdots, \mu_k)$ のまわりの標本の散乱行列とすれば，$\|v_{ij}\|$ の要素は次の p.d.f. $g(\{v_{ij}\})$ を持つ．すなわち $\|v_{ij}\|$ が正定値になる $\dfrac{1}{2}k(k+1)$ 次元の $\{v_{ij}\}$ の領域では

$$g(\{v_{ij}\}) = \frac{|\sigma^{ij}|^{\frac{1}{2}n}|v_{ij}|^{\frac{1}{2}(n-k-1)}\exp\left[-\frac{1}{2}\sum_{i,j=1}^{k}\sigma^{ij}v_{ij}\right]}{2^{\frac{1}{2}kn}\pi^{k(k-1)/4}\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-1}{2}\right)\cdots\Gamma\left(\frac{n-k+1}{2}\right)}. \tag{18.2.27}$$

その他の v に対しては

$$g(\{v_{ij}\}) = 0.$$

要素が p. d. f. (18.2.27) を持つ確率変数の行列式 $\|v_{ij}\|$ は，パラメータ k, n, $\|\sigma_{ij}\|$ で特徴づけられるウィシャート分布 $W(k, n, \|\sigma_{ij}\|)$ を持つといわれる．

Wishart (1928) は幾何学的方法で最初にこの分布を導いた．後に幾何学的な議論を若干修正して Mahalanobis, Bose と Roy (1937) によっても示された．この分布は他にいろいろな方法によっても導かれている．たとえば Ingham (1933), Wishart と Bartlett (1932), Madow (1938), Hsu (1939 *a*), Sverdrup (1947), Rasch (1948) による．ここで述べた導入方法は本質的にはイングハム，ウィシャートとバートレットが与えたものである．

$k=1$ のときは $\sigma^{11}=1/\sigma^2$ となる．ただし σ^2 は標本が抽出された正規分布の分散．したがってウィシャート分布 $W(1, n, \sigma^2)$ は $\sigma^{11}v_{11}$ が自由度 n のカイ 2 乗分布を持つような分布に帰着する．

標本における確率変数の個数は nk で，ウィシャート分布における確率変数の個数は $\frac{1}{2}k(k+1)$ である．これはウィシャート分布が標本要素の nk 次元分布から決まる $\frac{1}{2}k(k+1)$ 次元周辺分布であることを意味している．James (1954) は標本の確率素分を次の 3 つの独立な分布の確率素分に変換する $\{x_{i\xi};\, i=1,\cdots,k,\ \xi=1,\cdots,n\}$ の変換の存在を示している．独立な分布の 1 つはウィシャート分布で他の 1 つは k 個の n 次元ベクトル $[(x_{i1}-\mu_i),\cdots,(x_{in}-\mu_i)]$, $i=1,\cdots,k$ で張られた k 次元（超）平面に関する分布である．3 番目の分布はこの k 次元（超）平面における k 個のベクトルの方向を定める $k\times k$ の直交行列の分布である．

(b) 正規分布からの標本散乱のモーメントと分布

(18.1.26) では，標本 $(x_{1\xi},\cdots,x_{k\xi},\ \xi=1,\cdots,n)$ が抽出された分布の平均 $(\mu_1,\cdots,\mu_k)$ のまわりの標本の散乱 $|v_{ij}|$ の平均値が $k!\binom{n}{k}\cdot|\sigma_{ij}|$ であることをみた．ただし $\|\sigma_{ij}\|$ は分布の共分散行列である．特に k 次元正規分布の場合，$|v_{ij}|$ の r 次のモーメントは次のようにして求められる．

まず便宜上

$$g(\{v_{ij}\}) = K(k, n, \{\sigma_{ij}\})\cdot h(k, n, \{\sigma_{ij}\}, \{v_{ij}\}) \tag{18.2.28}$$

と置こう．ただし

$$K(k, n, \{\sigma_{ij}\}) = \frac{|\sigma^{ij}|^{\frac{1}{2}n}}{2^{\frac{1}{2}kn}\pi^{k(k-1)/4}\prod_{i=1}^{k}\Gamma\left(\frac{n+1-i}{2}\right)}. \tag{18.2.29}$$

そして $\|v_{ij}\|$ が正定値となる $\{v_{ij}\}$ の空間上で

$$h(k, n, \{\sigma_{ij}\}, \{v_{ij}\}) = |v_{ij}|^{\frac{1}{2}(n-k-1)}\exp\left[-\frac{1}{2}\sum_{i,j=1}^{k}\sigma^{ij}v_{ij}\right] \tag{18.2.30}$$

その他の $\{v_{ij}\}$ に対しては $h=0$ である．すると

$$\mathscr{E}(|v_{ij}|^r) = K(k, n, \{\sigma_{ij}\})\int_{R_{\frac{1}{2}k(k+1)}}|v_{ij}|^r h(k, n, \{\sigma_{ij}\}, \{v_{ij}\})\prod_{i\geq j=1}^{k}dv_{ij} \tag{18.2.31}$$

となる．

さて $g(\{v_{ij}\})$ の $\{v_{ij}\}$ の全標本空間上の積分は 1 だから

$$\int_{R_{\frac{1}{2}k(k+1)}}h(k, n, \{\sigma_{ij}\}, \{v_{ij}\})\prod_{i\geq j=1}^{k}dv_{ij} = \frac{1}{K(k, n, \{\sigma_{ij}\})} \tag{18.2.32}$$

が得られる．しかし (18.2.31) の被積分関数は $h(k, n+2r, \{\sigma_{ij}\}, \{v_{ij}\})$ で，この関数の $\{v_{ij}\}$ の標本空間にわたる積分は $1/K(k, n+2r, \{\sigma_{ij}\})$ である．よって

$$\mathscr{E}(|v_{ij}|^r) = \frac{K(k, n, \{\sigma_{ij}\})}{K(k, n+2r, \{\sigma_{ij}\})} = |2\sigma_{ij}|^r\prod_{i=1}^{k}\left[\frac{\Gamma\left(\frac{n+1-i}{2}+r\right)}{\Gamma\left(\frac{n+1-i}{2}\right)}\right]. \tag{18.2.33}$$

$r=1$ にすると $N(\mu_i, \|\sigma_{ij}\|)$ からの大きさ n の標本における $|v_{ij}|$ に対して

$$\mathscr{E}(|v_{ij}|) = k!\binom{n}{k}|\sigma_{ij}|$$

が見い出される．この結果は (18.1.26) からも導かれているが，もっと一般的に共分散行列 $\|\sigma_{ij}\|$ を持つ任意の k 次元分布からの大きさ n の標本における $|v_{ij}|$ に対しても成立する．

$\mathscr{E}(|v_{ij}|^r)$ は次のように書けることに注意しよう．

(18.2.34) $$\mathscr{E}(|v_{ij}|^r)=\int_0^\infty\cdots\int_0^\infty(|2\sigma_{ij}|z_1\cdots z_k)^r\prod_{i=1}^k\left[\frac{z_i^{\frac{1}{2}(n+1-i)-1}e^{-z_i}}{\Gamma\left(\frac{n+1-i}{2}\right)}dz_i\right].$$

したがって $|v_{ij}|$ の分布は，z_i をそれぞれガンマ分布 $G\left(\frac{1}{2}(n+1-i)\right)$，$i=1,\cdots,k$ を持つ独立な確率変数とするとき，$|2\sigma_{ij}|(z_1\cdots z_k)$ の分布に一致することがわかる．

要約すると次の結果を得る．

18.2.2　$|v_{ij}|$ が $N(\{\mu_i\};\|\sigma_{ij}\|)$ からの大きさ n の標本の母（集団）平均 $(\mu_1,\cdots,\mu_k)$ のまわりの散乱であれば，$|v_{ij}|$ の分布は $|2\sigma_{ij}|\prod_{i=1}^k z_i$ の分布に一致する．ただし z_i は独立な確率変数で，それぞれガンマ分布 $G\left(\frac{1}{2}(n+1-i)\right)$，$i=1,\cdots,k$ を持つ．

積分として表わされた $|v_{ij}|$ の分布関数の明確な表現に関しては Wilks (1932) を参照されたい．

(c)　ウィシャート分布の再生性

ウィシャート分布に対する再生性の法則は次のように述べられる．

18.2.3　$\{v_{ij}^{(1)}\}$，$\{v_{ij}^{(2)}\}$ はウィシャート分布 $W(k,n_1,\|\sigma_{ij}\|)$，$W(k,n_2,\|\sigma_{ij}\|)$ を持つ独立な確率変数の組とする．このとき確率変数の組 $\{v_{ij}^{(1)}+v_{ij}^{(2)}\}$ はウィシャート分布 $W(k,n_1+n_2,\|\sigma_{ij}\|)$ を持つ．

これは特性関数を用いれば，ただちに証明される．まず (18.2.4) から，$\{v_{ij}^{(1)}\}$，$\{v_{ij}^{(2)}\}$ の特性関数はそれぞれ

$$|\tau_{ij}|^{\frac{1}{2}n_1}\cdot|\tau_{ij}-\mathrm{i}\varepsilon_{ij}t_{ij}|^{-\frac{1}{2}n_1},\qquad|\tau_{ij}|^{\frac{1}{2}n_2}\cdot|\tau_{ij}-\mathrm{i}\varepsilon_{ij}t_{ij}|^{-\frac{1}{2}n_2}$$

である．ただし $\tau_{ij}=\frac{1}{2}\sigma^{ij}$ で $\varepsilon_{ii}=1$，$\varepsilon_{ij}=\frac{1}{2}$ $(i\neq j)$．$\{v_{ij}^{(1)}\}$ と $\{v_{ij}^{(2)}\}$ は独立だから $\{v_{ij}^{(1)}+u_{ij}^{(2)}\}$ の特性関数はこの 2 つの特性関数の積，すなわち

$$|\tau_{ij}|^{\frac{1}{2}(n_1+n_2)}\cdot|\tau_{ij}-\mathrm{i}\varepsilon_{ij}t_{ij}|^{-\frac{1}{2}(n_1+n_2)}$$

である．これはウィシャート分布 $W(k,n_1+n_2,\|\sigma_{ij}\|)$ の特性関数である．

ウィシャート分布や正規分布に対する再生性に類似した基本的な結果を次に述べよう．これはあとで役立つだろう．

18.2.4　$\{a_{ij}\}$ はウィシャート分布 $W(k,n,\|\sigma_{ij}\|)$ を持つ確率変数の組で $(b_{i\beta},\cdots,$

$b_{k\beta})$，$\beta=1,\cdots,p$ は p 個の独立な同一の分布 $N(\{0\};\|\sigma_{ij}\|)$ を持つ確率変数の組とする．しかも，これらの確率変数の組は $\{a_{ij}\}$ の組とも独立であるとする．このとき確率変数の組 $\{a_{ij}+\sum_{\beta=1}^p b_{i\beta}b_{j\beta}\}$ はウィシャート分布 $W(k,n+p,\|\sigma_{ij}\|)$ を持つ．

特性関数による **18.2.4** の証明は **18.2.3** と同様である．読者に問題として残しておこう．$p\geqq k$ に対しては **18.2.4** は **18.2.3** から直接に導かれる．

18.3　k 次元正規分布からの標本における平均と内部散乱行列との独立性

$(x_{1\xi},\cdots,x_{k\xi},\ \xi=1,\cdots,n)$ を k 次元分布 $N(\{\mu_i\};\|\sigma_{ij}\|)$ からの標本，$\|u_{ij}\|$ をこの標本の内部散乱行列とし，次を示そう．

18.3.1　内部散乱行列 $\|u_{ij}\|$ の要素と標本平均 $(\bar{x}_1,\cdots,\bar{x}_k)$ はそれぞれ分布 $W(k,n-1,\|\sigma_{ij}\|)$，$N(\{\mu_i\};\|\sigma_{ij}/n\|)$ を持つ確率変数の組で独立になる．

この結果を示すために

(18.3.1) $$\varphi(\{t_{ij}\},\{t_i\})=\mathscr{E}\left\{\exp\left[\mathrm{i}\sum_{i,j=1}^k v_{ij}t_{ij}+\mathrm{i}\sum_{i=1}^k(\bar{x}_i-\mu_i)t_i\right]\right\}$$

で定義された $\{v_{ii},\ i=1,\cdots,k;2v_{ij},\ i>j=1,\cdots,k\}$ と $((\bar{x}_1-\mu_1),\cdots,(\bar{x}_k-\mu_k))$ との特性関数を考える．ただし $\|v_{ij}\|$ は (18.1.27) で定義されたように母（集団）平均 $(\mu_1,\cdots,\mu_k)$ のまわりの標本の散乱行列である．(18.3.1) の右辺は

(18.3.2) $$\left(\mathscr{E}\left\{\exp\left[\mathrm{i}\sum_{i,j=1}^k y_iy_jt_{ij}+\mathrm{i}\sum_{i=1}^k\frac{y_i}{n}t_i\right]\right\}\right)^n$$

で表わされることに注意する．ただし $y_i=x_i-\mu_i$．(18.3.2) を 7.4 節の結果を用いて評価すれば

(18.3.3) $$\varphi(\{t_{ij}\},\{t_i\})=|\sigma^{ij}|^{\frac{1}{2}n}\cdot|\sigma_*^{ij}|^{-\frac{1}{2}n}\cdot\exp\left[-\frac{1}{2}\sum_{i,j=1}^k\sigma_{ij}^*\frac{t_it_j}{n}\right]$$

となる．ここに $\|\sigma_*^{ij}\|=\|\sigma^{ij}-2\mathrm{i}t_{ij}\|$，$\|\sigma_{ij}^*\|=\|\sigma_*^{ij}\|^{-1}$ である．

さて多次元のレヴィの定理を適用すれば，$\{v_{ii},(\bar{x}_i-\mu_i),\ i=1,\cdots,k;2v_{ij},\ i>j=$

$1, \cdots, k\}$ の p. d. f. は

$$(18.3.4) \qquad \left(\frac{1}{2\pi}\right)^{\frac{1}{2}k(k+3)} \int_{R^{\frac{1}{2}k(k+3)}} \exp\left[-\mathrm{i}\sum_{i,j=1}^{k} v_{ij}t_{ij} - \mathrm{i}\sum_{i=1}^{k}(\bar{x}_i - \mu_i)t_i\right] \cdot \varphi(\{t_{ij}\}, \{t_i\}) \prod_{i=1}^{k} dt_i \prod_{i \geqq j=1}^{k} dt_{ij}$$

で与えられる．7.4 節の結果を用いて t_i で積分すれば (18.3.4) は

$$(18.3.5) \qquad I \cdot \frac{\sqrt{|n\sigma^{ij}|}}{(2\pi)^{\frac{1}{2}k}} \exp\left[-\frac{n}{2}\sum_{i,j=1}^{k} \sigma^{ij}(\bar{x}_i - \mu_i)(\bar{x}_j - \mu_j)\right]$$

に帰着する．ただし

$$(18.3.6) \qquad I = \left(\frac{1}{2\pi}\right)^{\frac{1}{2}k(k+1)} \int_{R^{\frac{1}{2}k(k+1)}} |\sigma^{ij}|^{\frac{1}{2}(n-1)} |\sigma^{ij} - 2\mathrm{i}t_{ij}|^{-\frac{1}{2}(n-1)} \cdot \exp\left[-\mathrm{i}\sum_{i,j=1}^{k} u_{ij}t_{ij}\right] \sum_{i \geqq j=1}^{k} dt_{ij}$$

で

$$u_{ij} = v_{ij} - n(\bar{x}_i - \mu_i)(\bar{x}_j - \mu_j) = \sum_{\xi=1}^{n} (x_{i\xi} - \bar{x}_i)(x_{j\xi} - \bar{x}_j).$$

よって (18.3.6) と (18.2.6) を比較すれば I は確率変数の組 (u_{ii}, $i = 1, \cdots, k$; $2u_{ij}$, $i > j = 1, \cdots, k$) の p. d. f. になる．ただし $\{u_{ij}\}$ はウィシャート分布 $W(k, n-1, \|\sigma_{ij}\|)$ を持つ．

このように (18.3.5) はウィシャート分布 $W(k, n-1, \|\sigma_{ij}\|)$ の p. d. f. と正規分布 $N\left(\{\mu_i\} ; \left\|\frac{\sigma_{ij}}{n}\right\|\right)$ の p. d. f. との積に書ける．すなわち **18.3.1** が示された．

18.4　ホテリングの一般化スチューデント分布

(a)　1 標本の場合

k 次元正規分布からの標本がある．いま，この分布の平均値ベクトルが，与えられたある定数ベクトル ($\mu_1^\circ, \cdots, \mu_k^\circ$) であるという仮説を検定したい．散乱による解析に基づく検定は標本の内部散乱 $|u_{ij}|$ と特定の平均値 ($\mu_1^\circ, \cdots, \mu_k^\circ$) のまわりの標本の散乱 $|v_{ij}|$ とを比較することである．これは比

$$(18.4.1) \qquad R_1 = \frac{|u_{ij}|}{|v_{ij}|} = \frac{|u_{ij}|}{|u_{ij} + n(\bar{x}_i - \mu_i^\circ)(\bar{x}_j - \mu_j^\circ)|}$$

をとることを暗示している．しかし

$$(18.4.2) \qquad |v_{ij}| = |u_{ij}| \cdot \left[1 + n\sum_{i,j=1}^{k} u^{ij}(\bar{x}_i - \mu_i^\circ)(\bar{x}_j - \mu_j^\circ)\right]$$

だから R_1 は

$$(18.4.3) \qquad R_1 = \frac{1}{1 + nD^2/(n-1)}$$

と書き改められる．ただし

$$(18.4.4) \qquad D^2 = (n-1)\sum_{i,j=1}^{k} u^{ij}(\bar{x}_i - \mu_i^\circ)(\bar{x}_j - \mu_j^\circ)$$

で $\|u^{ij}\| = \|u_{ij}\|^{-1}$.

明らかに R_1 は $[0, 1]$ 上にある．R_1 が 1 であるための必要十分条件は標本平均 ($\bar{x}_1, \cdots, \bar{x}_k$) がベクトルとして母 (集団) 平均 ($\mu_1^\circ, \cdots, \mu_k^\circ$) に等しいことである．もちろん $\|u_{ij}\|$ は正定値であると仮定する．$n > k$ ならば確率 1 で正定値になる．量 D^2 は標本と母集団間の **Mahalanobis** (1936) の (平方) 距離として知られている．あるいは簡単にマハラノビスの D^2 と呼ばれている．もちろん D^2 の値が大きいほど，R_1 の値は小さくなる．さらに

$$(18.4.5) \qquad T^2 = nD^2$$

で定義された量 T^2 を **Hotelling** (1931) の (平方) 一般化スチューデント比，あるいは簡単にホテリングの T^2 という．

Ω を $\|\sigma_{ij}\|$ が正定値となる σ_{ij}, $i \geqq j = 1, \cdots, k$ のすべての値および μ_1 のすべての値からなるパラメータ空間における点の許容集合とし，ω を $\mu_i = \mu_i^\circ$, $i = 1, \cdots, k$ となるパラメータ空間の点からなる Ω の部分集合とする．このとき λ を複合仮説 $\mathscr{H}(\omega ; \Omega)$ の検定 (13.3 節を見よ) におけるネイマン=ピヤソン尤度比とすれば，$R_1 = \lambda^{2/n}$ であることに注意しよう．証明は読者自身で試みよ．

さて $\mathscr{H}(\omega ; \Omega)$ が正しい，すなわち標本が抽出される正規分布が平均値ベクトル ($\mu_1^\circ, \cdots, \mu_k^\circ$) を持つとき，$R_1$ の標本分布を考察しよう．この問題に関する比較的簡単な方法は，まず R_1 のモーメントを計算してモーメントから分布を推測する方法である．したがって

$$\mathscr{E}(R_1^r) = \mathscr{E}(|u_{ij}|^r |v_{ij}|^{-r}), \qquad r = 1, 2, \cdots$$

の値を求める．

$\|v_{ij}\|$ の要素はウィシャート分布 $W(k, n, \|\sigma_{ij}\|)$ を持つので，$\mathscr{E}(|v_{ij}|^{-r})$ は (18.2.33) において r を $-r$ に置き換えることによって与えられる．しかし $v_{ij} = u_{ij} + b_i b_j$．ただし $b_i = \sqrt{n}(\bar{x}_i - \mu_i^0)$ で $\{u_{ij}\}$ と $\{b_i\}$ は **18.3.1** で示されたように，それぞれ分布 $W(k, n-1, \|\sigma_{ij}\|)$，$N(\{0\}; \|\sigma_{ij}\|)$ に従う確率変数の組である．

したがって次のように書ける．

$$(18.4.6) \qquad K(k, n-1, \{\sigma_{ij}\})\left[\frac{|\sigma^{ij}|}{(2\pi)^k}\right]^{\frac{1}{2}} \int_{R^{\frac{1}{2}k(k+1)}} \int_{R^k} |v_{ij}|^{-r} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{k} \sigma^{ij} b_i b_j\right]$$

$$\cdot\, h(k, n-1, \{\sigma_{ij}\}, \{u_{ij}\}) \prod_{i=1}^{k} db_i \prod_{i \geq j=1}^{k} du_{ij}$$

$$= |2\sigma_{ij}|^{-r} \prod_{i=1}^{k} \frac{\Gamma\left(\frac{n+1-i}{2} - r\right)}{\Gamma\left(\frac{n+1-i}{2}\right)}.$$

ただし $K(k, n-1, \{\sigma_{ij}\})$ と $h(k, n-1, \{\sigma_{ij}\}, \{u_{ij}\})$ は 18.2(b) 節で定義されている．

(18.4.6) において n を $n+2r$ で置き換えた式の両辺に $K(k, n-1, \{\sigma_{ij}\})/K(k, n+2r-1, \{\sigma_{ij}\})$ を乗じると左辺は $\mathscr{E}(R_1^r)$ を定義する積分になり，右辺はこの積分値になる．したがって次式を得る．

$$(18.4.7) \qquad \mathscr{E}(R_1^r) = \left[\frac{|2\sigma_{ij}|^{-r} K(k, n-1, \{\sigma_{ij}\})}{K(k, n+2r-1, \{\sigma_{ij}\})}\right] \prod_{i=1}^{k} \frac{\Gamma\left(\frac{n+1-i}{2}\right)}{\Gamma\left(\frac{n+1-i}{2} + r\right)}.$$

(18.4.7) の $K(k, n-1, \{\sigma_{ij}\})$，$K(k, n+2r-1, \{\sigma_{ij}\})$ に (18.2.29) の値を代入して簡単な計算を行なえば

$$(18.4.8) \qquad \mathscr{E}(R_1^r) = \frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{n-k}{2} + r\right)}{\Gamma\left(\frac{n}{2} + r\right)\Gamma\left(\frac{n-k}{2}\right)}$$

が $r = 1, 2, \cdots$ について成立する．(18.4.8) の右辺は次のようにベータ分布の r 次のモーメントとして表現されることがわかる．

$$\mathscr{E}(R_1^r) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{n-k}{2}\right)} \int_0^1 R_1^{\frac{1}{2}(n-k)-1+r}(1-R_1)^{\frac{1}{2}k-1}\, dR_1, \qquad r = 0, 1, 2, \cdots.$$

5.5.1a により R_1 はベータ分布 $Be\left(\frac{1}{2}(n-k), \frac{1}{2}k\right)$ を持つ，すなわち R_1 の確率素分は

$$(18.4.9) \qquad \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{n-k}{2}\right)} R_1^{\frac{1}{2}(n-k)-1}(1-R_1)^{\frac{1}{2}k-1}\, dR_1$$

であることがわかる．

要約すると次のようになる．

18.4.1 $(x_{1\xi}, \cdots, x_{k\xi},\ \xi = 1, \cdots, n)$，$n > k$ が正規分布 $N(\{\mu_i^0\}; \|\sigma_{ij}\|)$ からの標本ならば散乱の比

$$R_1 = |u_{ij}|/|v_{ij}|$$

はベータ分布 $Be\left(\frac{1}{2}(n-k), \frac{1}{2}k\right)$ を持つ．

(18.4.9) に変換

$$R_1 = \frac{1}{1 + T^2/(n-1)}$$

を適用して T^2 の標本空間が $(0, +\infty)$ であることを考慮すれば，ホテリングの T^2 の確率素分として

$$(18.4.10) \qquad \frac{\Gamma\left(\frac{n}{2}\right)}{(n-1)^{\frac{1}{2}k}\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{n-k}{2}\right)}(1 + T^2/(n-1))^{-\frac{1}{2}n}(T^2)^{\frac{1}{2}k-1}\, d(T^2)$$

を得る．特に $k = 1$ ならば (18.4.10) は自由度 $(n-1)$ のスチューデント t^2 の確率素分に帰着する．[7.8(b) 節を参照]．

(b) 2標本の場合

2つの標本 $(x_{1\xi_1}^{(1)}, \cdots, x_{k\xi_1}^{(1)},\ \xi_1 = 1, \cdots, n_1)$，$(x_{1\xi_2}^{(2)}, \cdots, x_{k\xi_2}^{(2)},\ \xi_2 = 1, \cdots, n_2)$ はともに同一のパラメータの集合を持つ正規分布からの標本とする．この共通の分布を$N(\{\mu_i\}; \|\sigma_{ij}\|)$で表わす．$\|u_{ij}^{(1)}\|$，$\|u_{ij}^{(2)}\|$ はそれぞれの標本の内部散乱行列とし，$\|u_{ij}\|$ は2つの標本を1つに合わせた標本の内部散乱行列とする．行列 $\|u_{ij}^{(1)} + u_{ij}^{(2)}\|$ を2つの標本に対する標本内散乱行列という．幾何学的に述べれば標本内散乱行列は次のようにして得られるk次元クラスターに対する散乱行列である．すなわち一方の標本クラスターを両方の標本平均が一致するまで他方の標本クラスター上に移動させて（回転することなく），2つの標本

クラスターを 1 つの標本クラスターとして合わせることである．

比

$$R_1' = \frac{|u_{ij}^{(1)} + u_{ij}^{(2)}|}{|u_{ij}|} \tag{18.4.11}$$

を考察しよう．R_1' は (18.4.3) で定義された R_1 と同じ性質と分布を持つことがわかるだろう．まず

$$u_{ij}^W = u_{ij}^{(1)} + u_{ij}^{(2)} \tag{18.4.12}$$

と置こう．u_{ij} を (18.1.20) の右辺の代表元とすれば，u_{ij} は次のように表わされる．

$$u_{ij} = u_{ij}^W + b_i b_j. \tag{18.4.13}$$

ただし

$$b_i = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}(\bar{x}_i^{(1)} - \bar{x}_i^{(2)}). \tag{18.4.14}$$

したがって

$$R_1' = \frac{|u_{ij}^W|}{|u_{ij}^W + b_i b_j|} \tag{18.4.15}$$

と書ける．

18.2.3 より $\{u_{ij}^W\}$ はウィシャート分布 $W(k, n_1 + n_2 - 2, \|\sigma_{ij}\|)$ を持つ．また b_i は 2 つの標本平均の関数だから，**18.3.1** が使えて $\{b_i\}$ は $\{u_{ij}^W\}$ と独立になる．事実，読者は $\{b_i\}$ が正規分布 $N(\{0\}; \|\sigma_{ij}\|)$ に従うことを容易に確かめることができよう．

このように R_1' の分布を求める問題は (18.4.1) における R_1 の分布を見い出す問題と同じになる．事実，R_1' の分布は R_1 の分布において $n-1$ を $n_1 + n_2 - 2$ で置き換えた分布に一致する．

また R_1' は次のように書ける．

$$R_1' = \frac{1}{1 + \dfrac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D'^2}. \tag{18.4.16}$$

ただし

$$\begin{aligned} D'^2 &= \frac{(n_1 + n_2)(n_1 + n_2 - 2)}{n_1 n_2} \sum_{i,j=1}^{k} u_W^{ij} b_i b_j \\ &= (n_1 + n_2 - 2) \sum_{i,j=1}^{k} u_W^{ij} (\bar{x}_i^{(1)} - \bar{x}_i^{(2)})(\bar{x}_j^{(1)} - \bar{x}_j^{(2)}) \end{aligned} \tag{18.4.17}$$

で $\|u_W^{ij}\| = \|u_{ij}^W\|^{-1}$．

D'^2 は Mahalanobis (1930, 1936) の 2 標本間の（平方）一般化距離である．

2 標本問題に対するホテリングの一般化スチューデント比 T'^2 は関係式

$$T'^2 = \frac{n_1 n_2}{n_1 + n_2} D'^2$$

で定義される．T'^2 の分布の確率素分は (18.4.10) の $n-1$ を $n_1 + n_2 - 2$ で置き換えたものである．もちろん，ここでは 2 つの標本は同一の k 次元正規分布から独立に抽出されると仮定している．

最後に $R_1' = \lambda^{2/n_1+n_2}$ に注意しよう．ただし λ は $\mathscr{H}(\omega; \Omega)$ の仮説検定に対するネイマン＝ピヤソンの尤度比である．ここでは Ω は $\|\sigma_{ij}^{(1)}\| = \|\sigma_{ij}^{(2)}\| = \|\sigma_{ij}\|$ が正定値となり平均値ベクトル $(\mu_1^{(1)}, \cdots, \mu_k^{(1)})$，$(\mu_1^{(2)}, \cdots, \mu_k^{(2)})$ が実数となるようなパラメータ空間で，ω は 2 つの平均値ベクトルが等しくなる Ω の部分空間である．ただし $N(\{\mu_i^{(1)}\}; \|\sigma_{ij}^{(1)}\|)$ と $N(\{\mu_i^{(2)}\}; \|\sigma_{ij}^{(2)}\|)$ は 2 つの標本が抽出される正規分布を表わしている．この確認は読者にまかそう．

18.5　多次元モデルI分散分析検定

一般型における 1 次元モデル I 分散分析検定は 13.3(b) 節で定義された正規回帰論に関する線形仮説 $\mathscr{H}$ の統計的仮説検定である．この仮説に対する尤度比 λ は次の型で表わされる［(13.3.9) を見よ］．

$$\lambda = \left[\frac{S_\Omega}{S_\Omega + (S_\omega - S_\Omega)}\right]^{\frac{1}{2}n}. \tag{18.5.1}$$

ただし S_Ω，$S_\omega - S_\Omega$ は $\mathscr{H}$ が正しいとき，独立な確率変数であって，それぞれ 1 次元ウィシャート分布 $W(1, m_1, \sigma^2)$，$W(1, m_2, \sigma^2)$ を持つ．ただし $m_1 = n - k$，$m_2 = k - k'$ はそれぞれ S_Ω，$S_\omega - S_\Omega$ における自由度．13.3(b) 節で指摘されたように尤度比

λ はスネディッカーの $\mathscr{F}$ 検定に同値である．ただし

$$\mathscr{F}=\frac{m_1(S_\omega-S_\Omega)}{m_2S_\Omega}. \tag{18.5.2}$$

$\mathscr{H}$ の検定に関しては λ は $\lambda^{2/n}$ と同値である．もちろん $\lambda^{2/n}$ は (18.5.1) における［ ］内の比である． 一般型において， モデルＩ分散分析検定を多次元へ拡張するためには，(18.5.1) における［ ］内の比を一般化すればよい．この一般化は

$$R_s=\frac{|a_{ij}|}{\left|a_{ij}+\sum_{\beta=1}^{s}b_{i\beta}b_{j\beta}\right|},\qquad i,j=1,\cdots,k,\ \ k\leqslant m \tag{18.5.3}$$

なる型の比である．ただし $\{a_{ij}\}$ はウィシャート分布 $W(k,m,\|\sigma_{ij}\|)$ を持ち，$(b_{1\beta},\cdots,b_{k\beta})$，$\beta=1,\cdots,s$ はすべて正規分布 $N(\{0\};\|\sigma_{ij}\|)$ に従うベクトル確率変数である． さらにすべての $(b_{1\beta},\cdots,b_{k\beta})$，$\beta=1,\cdots,s$ は $\{a_{ij}\}$ と独立である．

R_s の r 次のモーメントを計算するために，R_1 の r 次のモーメントを求めるとき用いた方法を使えば，(18.4.8) で与えられたように

$$\mathscr{E}(R_s^r)=\prod_{i=1}^{k}\frac{\Gamma\left(\frac{m+1-i}{2}+r\right)\Gamma\left(\frac{m+s+1-i}{2}\right)}{\Gamma\left(\frac{m+1-i}{2}\right)\Gamma\left(\frac{m+s+1-i}{2}+r\right)} \tag{18.5.4}$$

が得られる．

$s\geqslant k$ ならば，$\mathscr{E}(R_s^r)$ は次のように表わされる．

$$\mathscr{E}(R_s^r)=\prod_{i=1}^{k}\left[\frac{\Gamma\left(\frac{m+s+1-i}{2}\right)}{\Gamma\left(\frac{m+1-i}{2}\right)\Gamma\left(\frac{s}{2}\right)}\cdot\frac{\Gamma\left(\frac{m+1-i}{2}+r\right)\Gamma\left(\frac{s}{2}\right)}{\Gamma\left(\frac{m+s+1-i}{2}+r\right)}\right]. \tag{18.5.4 a}$$

よって z_i を独立で，それぞれベータ分布

$$Be\left(\frac{m+1-i}{2},\frac{s}{2}\right),\qquad i=1,\cdots,k$$

に従う確率変数とすれば，R_s の分布は $\prod_{i=1}^{k}z_i$ の分布に一致することがわかる．

同様に $1\leqslant s<k$ ならば，$\mathscr{E}(R_s^r)$ は次の型に書ける．

$$\mathscr{E}(R_s^r)=\prod_{i=1}^{s}\left[\frac{\Gamma\left(\frac{m+i}{2}\right)}{\Gamma\left(\frac{m-k+i}{2}\right)\Gamma\left(\frac{k}{2}\right)}\cdot\frac{\Gamma\left(\frac{m-k+i}{2}+r\right)\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{m+i}{2}+r\right)}\right]. \tag{18.5.4 b}$$

すなわち R_s の分布は積 $\prod_{i=1}^{s}w_i$ の分布に一致する． ただし w_i は独立な確率変数で，

それぞれベータ分布

$$Be\left(\frac{m-k+i}{2},\frac{k}{2}\right),\qquad i=1,\cdots,s$$

を持つ．

要約すると次の結果が得られる．

18.5.1　$\|a_{ij}\|$ は対称行列で，確率 1 で正定値とする．$\|a_{ij}\|$ の要素はウィシャート分布 $W(k,m,\|\sigma_{ij}\|)$ を持ち，$(b_{1\beta},\cdots,b_{k\beta})$，$\beta=1,\cdots,s$ は独立ですべて正規分布 $N(\{0\};\|\sigma_{ij}\|)$ を持つ s 個の k 次元確率変数とする．さらに $\|a_{ij}\|$ における確率変数とも独立であるとする．いま

$$R_s=\frac{|a_{ij}|}{\left|a_{ij}+\sum_{\beta=1}^{s}b_{i\beta}b_{j\beta}\right|}$$

とすれば

（i）$s\geqslant k$ のとき R_s の分布は独立で，それぞれベータ分布

$$Be\left(\frac{m+1-i}{2},\frac{s}{2}\right),\qquad i=1,\cdots,k$$

を持つ k 個の確率変数の積の分布に一致する．

（ii）$1\leqslant s<k$ のとき R_s の分布は独立で，それぞれベータ分布

$$Be\left(\frac{m-k+i}{2},\frac{k}{2}\right),\qquad i=1,\cdots,s$$

を持つ s 個の確率変数の積の分布に一致する．

(18.4.1) で定義された R_1 は，R_s において $s=1$，$m=n-1$，$a_{ij}=u_{ij}$，$b_i=\sqrt{n}(\bar{x}_i-\mu_i)$，$i,j=1,\cdots,k$ とした特別の場合であることがわかる．同様に (18.4.11) あるいは (18.4.15) で定義された R_1' は R_s において $s=1$，$m=n_1+n_2-2$，$a_{ij}=u_{ij}^{(1)}+u_{ij}^{(2)}$，

$$b_i=\sqrt{\frac{n_1n_2}{n_1+n_2}}(\bar{x}_i^{(1)}-\bar{x}_i^{(2)}),\qquad i,j=1,\cdots,k$$

とした特別の場合である．$s=2$ に対しては，R_s の興味ある特殊な場合が起る．このとき $\sqrt{R_2}$ はベータ分布 $Be(m+1-k,k)$ を持つ．すなわち $\sqrt{R_2}$ の確率素分は

$$\frac{\Gamma(m+1-k)}{\Gamma(m+1)\Gamma(k)}(\sqrt{R_2})^{m-k}(1-\sqrt{R_2})^{k-1}d\sqrt{R_2} \tag{18.5.5}$$

になる．

これを確かめるには R_2 のモーメントを求めればよい．(18.5.4 b) より，これは

$$\mathscr{E}(R_2^r) = \frac{\Gamma\left(\frac{m+2}{2}\right)\Gamma\left(\frac{m+1}{2}\right)\Gamma\left(\frac{m+2-k}{2}+r\right)\Gamma\left(\frac{m+1-k}{2}+r\right)}{\Gamma\left(\frac{m+2}{2}+r\right)\Gamma\left(\frac{m+1}{2}+r\right)\Gamma\left(\frac{m+2-k}{2}\right)\Gamma\left(\frac{m+1-k}{2}\right)} \tag{18.5.6}$$

となる．ルジャンドルの重複公式 (7.6.15)，すなわち

$$\Gamma(\nu)\Gamma\left(\nu+\frac{1}{2}\right) = \frac{\sqrt{\pi}\,\Gamma(2\nu)}{2^{2\nu-1}} \tag{18.5.7}$$

を用いると (18.5.6) のガンマ関数は 4 つに縮められる．すなわち

$$\mathscr{E}(R_2^r) = \frac{\Gamma(m+1)\Gamma(m+1-k+2r)}{\Gamma(m+1+2r)\Gamma(m+1-k)} \tag{18.5.8}$$

を得る．これは次のように書ける．

$$\mathscr{E}(R_2^r) = \frac{\Gamma(m+1)}{\Gamma(m+1-k)\Gamma(k)}\int_0^1 z^{m-k+2r}(1-z)^{k-1}\,dz \tag{18.5.9}$$

$$r = 1, 2, \cdots.$$

(18.5.9) および **5.5.1 a** より，z をベータ分布 $Be(m+1-k, k)$ に従うとすれば，R_2 の分布は z^2 の分布に一致する．　よって $\sqrt{R_2}$ はこのベータ分布に従う．　すなわち (18.5.5) の確率素分を持つ．　R_s の分布の積分型による表現は Wilks (1932) で与えられている．

例題　$(x_{1\xi_\gamma}^{(\gamma)}, \cdots, x_{k\xi_\gamma}^{(\gamma)},\ \xi_\gamma = 1, \cdots, n_\gamma)$，$\gamma = 1, 2, 3$ は同一の正規分布 $N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ からの 3 つの独立な標本とする．$\|u_{ij}^{(\gamma)}\|$，$\gamma = 1, 2, 3$ は 3 つの標本の内部散乱行列で，$\|u_{ij}\|$ は 3 つの標本を 1 つに合わせた(大)標本の内部散乱行列とする．$\|u_{ij}^{(1)} + u_{ij}^{(2)} + u_{ij}^{(3)}\| = \|u_{ij}^W\|$ を 3 標本に対する標本内散乱行列とする．このとき比

$$R_2' = \frac{|u_{ij}^W|}{|u_{ij}|} \tag{18.5.10}$$

は (18.4.15) における比 R_1' の 3 標本への拡張である．R_2' は (18.5.3) の R_s において $s=2$，$m = n_1 + n_2 + n_3 - 3$ とした特別の場合である．　したがって $\sqrt{R_2'}$ の確率素分は (18.5.5) の m を $n_1 + n_2 + n_3 - 3$ に置き換えることによって表わされる．

確率変数の組 $\{u_{ij}^W\}$ はウィシャート分布 $W(k, n_1 + n_2 + n_3 - 3, \|\sigma_{ij}\|)$ を持つ．さらに標本間散乱行列 $\|u_{ij}^B\|$ は

$$u_{ij}^B = u_{ij} - u_{ij}^W = \sum_{\gamma=1}^{3} n_\gamma(\bar{x}_i^{(\gamma)} - \bar{x}_i)(\bar{x}_j^{(\gamma)} - \bar{x}_j), \qquad i, j = 1, \cdots, k \tag{18.5.11}$$

で定義される．ただし $(\bar{x}_1, \cdots, \bar{x}_k)$ は 3 つを合わせた標本の平均である．次の 3 つの条件を満たす 2 つの独立なベクトル $(b_{1\beta}, \cdots, b_{k\beta})$，$\beta = 1, 2$ が見い出される．（i）$(b_{1\beta}, \cdots, b_{k\beta})$，$\beta = 1, 2$ は $\{u_{ij}^W\}$ と独立で，（ii）$b_{i\beta} = \sum_{\gamma=1}^{3} c_{i\beta}^{(\gamma)}\bar{x}_i^{(\gamma)}$ で，各ベクトル（の要素）は分布 $N(\{0\}\,;\|\sigma_{ij}\|)$ を持ち，(iii) (18.5.11) の右辺が $\sum_{\beta=1}^{2} b_{i\beta}b_{j\beta}$ になる．しかし R_2' の分布を求めるには，これらのベクトルを定める必要はない．この分布は R_1 のときと同様にモーメントを計算する方法で見い出される．

もちろん R_2' は 18.4 (b) 節の終わりで示されたネイマン＝ピヤソンの尤度比を 3 標本に拡張したときの尤度比と同値であることも示される．

18.6　主　成　因

(a)　散乱行列の固有値と固有ベクトル

本節では，k 次元分布からの大きさ n $(k \leqslant n)$ の標本を考える．この標本は幾何学的には k 次元ユークリッド空間 R_k 内の n 個の点からなる標本クラスターとして表わされる．この標本クラスターを s 次元 $(s \leqslant k)$ ユークリッド空間 R_s 上へ直交射影して，射影された標本の s 次元散乱が最大になるようにしたい．　問題は （i）R_k の座標系に関する射影の方向と （ii）R_s における散乱の大きさを求めることである．

この統計的な問題の解は Hotelling (1933) による次の結果に述べられている．

18.6.1　$(x_{1\xi}, \cdots, x_{k\xi},\ \xi = 1, \cdots, n)$ は共分散行列が正定値の k 次元分布からの大きさ $n > k$ の標本とする．$\|u_{ij}\|$ はこの標本の内部散乱行列で，確率 1 で正定値であるとする．

$$(c_{1p}, \cdots, c_{kp}), \qquad p = 1, \cdots, s \tag{18.6.1}$$

は長さ 1 の k 次元ベクトル，すなわち $\sum_{i=1}^{k} c_{ip}^2 = 1$，$p = 1, \cdots, s$ とする．そして

$$z_{p\xi} = \sum_{i=1}^{k} c_{ip}x_{i\xi}, \quad p = 1, \cdots, s \tag{18.6.2}$$

と置く．標本 $(z_{1\xi},\cdots,z_{s\xi},\ \xi=1,\cdots,n)$ の内部散乱行列を $\|\bar{u}_{pq}\|$ とする．

散乱 $|\bar{u}_{pq}|$ を最大にするベクトル (18.6.1) の値は方程式系

(18.6.3) $$\sum_{j=1}^{k}(u_{ij}-l_p\delta_{ij})c_{jp}=0, \qquad i=1,\cdots,k,\ \ p=1,\cdots,s$$

の解である．ただし $l_1,\cdots,l_s$ は固有方程式

$$|u_{ij}-l\delta_{ij}|=0$$

の大きい方から s 番目までの根である．ここに δ_{ij} はクロネッカーの δ で，確率 1 で $l_s<\cdots<l_1$ であるとする．さらに，このベクトルは直交していて，$|\bar{u}_{pq}|$ の最大値は積 $l_1l_2\cdots l_s$ になる．

18.6.1 を示すには次のような標本の各要素の成分の一般的な線形関数

(18.6.4) $$z_\xi=\sum_{i=1}^{k}c_ix_{i\xi}, \qquad \xi=1,\cdots,n$$

を考える．そして平方和

(18.6.5) $$Q=\sum_{\xi=1}^{n}(z_\xi-\bar{z})^2=\sum_{i,j=1}^{k}u_{ij}c_ic_j$$

をとる．ただし $\sum_{i=1}^{k}c_i^2=1$.

さて Q が定留関数になる $(c_1,\cdots,c_k)$ の値を求めよう．ラグランジュ乗数の方法を用いれば，方程式

(18.6.6) $$\frac{\partial\varphi}{\partial c_i}=0, \qquad i=1,\cdots,k$$

から求めるベクトルが得られる．ただし

$$\varphi=Q+l\left(1-\sum_{1}^{k}c_i^2\right).$$

方程式 (18.6.6) は

(18.6.7) $$\sum_{j=1}^{k}(u_{ij}-l\delta_{ij})c_j=0, \qquad i=1,\cdots,k$$

になる．

(18.6.7) の自明な解，すなわち $c_i=0,\ i=1,\cdots,k$ はもちろん条件 $\sum_{i=1}^{k}c_i^2=1$ に反している．(18.6.7) の他の解を求めるために

(18.6.8) $$|u_{ij}-l\delta_{ij}|=0$$

を解く必要がある．

大きさ n $(>k)$ の標本は k 次元分布から抽出され，$\|u_{ij}\|$ は確率 1 で正定値を仮定しているので (18.6.8) は確率 1 で，k 個の異なる正の根を持つ．すなわち確率 1 で根を $0<l_k<\cdots<l_1$ としてよい．この根を行列 $\|u_{ij}\|$ の固有値という．また行列 $\|u_{ij}\|$ の特性値，特性根とも呼ばれている．

この根の任意の 1 つを $l_{p'}$ とすれば $(c_{1p'},\cdots,c_{kp'})$ は p' 番目の固有ベクトルであり

(18.6.9) $$\sum_{j=1}^{k}(u_{ij}-l_{p'}\delta_{ij})c_{jp'}=0$$

を満足する．

(18.6.9) 式に $c_{ip'}$ を乗じて i で和をとれば

(18.6.10) $$\sum_{i,j=1}^{k}(u_{ij}-l_{p'}\delta_{ij})c_{ip'}c_{jp'}=0$$

を得る．よって

(18.6.11) $$l_{p'}=\sum_{i,j=1}^{k}u_{ij}c_{ip'}c_{jp'}$$

を得る．

同様に，任意の異なる 2 根 $l_{p'}$, $l_{q'}$, $p'\neq q'$ に対して

(18.6.12) $$\sum_{i,j=1}^{k}(u_{ij}-l_{p'}\delta_{ij})c_{ip'}c_{jq'}=0$$
$$\sum_{i,j=1}^{k}(u_{ij}-l_{q'}\delta_{ij})c_{iq'}c_{jp'}=0$$

となるから（後式から前式を引くと）

(18.6.13) $$(l_{p'}-l_{q'})\sum_{i=1}^{k}c_{ip'}c_{iq'}=0$$

を得る．ゆえに $p'\neq q'$ に対して

(18.6.14) $$\sum_{i=1}^{k}c_{ip'}c_{iq'}=0.$$

すなわち固有ベクトル $(c_{1p},\cdots,c_{kp})$, $p=1,\cdots,k$ は互いに直交する．このベクトルは $\|u_{ij}\|$ の特性ベクトルあるいは特有ベクトルとも呼ばれている．

方程式 (18.6.12) のどちらかに (18.6.14) の結果を用いれば，$p' \neq q'$ に対して

$$\sum_{i,j=1}^{k} u_{ij} c_{ip'} c_{jq'} = 0 \tag{18.6.15}$$

を得る．

さて s $(s \leqq k)$ 番目までの根 $l_s < \cdots < l_1$ に対応する s 個の固有ベクトルを選び，

$$z_{p\xi} = \sum_{i=1}^{k} c_{ip} x_{i\xi}, \qquad p = 1, \cdots, s$$

とし標本 $(z_{1\xi}, \cdots, z_{s\xi},\ \xi = 1, \cdots, n)$ を構成したとする．長さ 1 のベクトルの直交性より，この標本はもとの標本を s 次元ユークリッド空間上へ直交射影した標本になっている．よって射影された標本の内部散乱は $|\bar{u}_{pq}|$, $p, q = 1, \cdots, s$ である．ただし

$$\bar{u}_{pq} = \sum_{\xi=1}^{n} (z_{p\xi} - \bar{z}_p)(z_{q\xi} - \bar{z}_q) = \sum_{i,j=1}^{k} u_{ij} c_{ip} c_{jq}. \tag{18.6.16}$$

しかし (18.6.11), (18.6.15) を用いれば

$$|\bar{u}_{pq}| = l_1 l_2 \cdots l_s \tag{18.6.17}$$

になる．これで **18.6.1** の証明を終わる．

$s = k$ にとれば，$|\bar{u}_{pq}| = |u_{ij}| = l_1 l_2 \cdots l_k$ がわかる．これは (18.6.8) の根の積が値 $|u_{ij}|$ を持つことからも明らかである．

ふたたび (18.6.8) に戻ると次がわかる．

$$l_1 + \cdots + l_k = \sum_{i=1}^{k} u_{ii}. \tag{18.6.18}$$

言い換えれば $\|u_{ij}\|$ の固有値の和は標本の個々の成分自身の平均のまわりの散乱の和に等しい．標本の内部散乱行列 $\|u_{ij}\|$ の固有値 $l_1, l_2, \cdots, l_k$ のおのおのを $n-1$ で割った値は Hotelling (1933) によって全標本分散の**主成因**と呼ばれている．なぜならば (18.6.18) からわかるように

$$\frac{l_1}{n-1} + \cdots + \frac{l_k}{n-1} = s_1^2 + \cdots + s_k^2 \tag{18.6.19}$$

である．ただし $s_i^2 = u_{ii}/(n-1)$.

(b) 散乱行列の固有値の標本論

$\|u_{ij}\|$ を標本の内部散乱行列とするとき固有値 $l_1, \cdots, l_k$ は行列方程式 (18.6.8) の根

だから固有値自身も確率変数である．もちろんこれらの根の標本論がどうなるだろうかという問題が考えられる．固有根の分布を正確に定める問題は本質的には標本 $(x_{1\xi}, \cdots, x_{k\xi},\ \xi = 1, \cdots, n)$ が k 次元球型正規分布から抽出された場合，すなわち共分散行列 $\|\sigma_{ij}\|$ がすべて等しい固有値を持つ場合だけ解決されている．本節ではこの標本論を述べよう．この問題は最初，近似的に Fisher (1939), Girshick (1939), Hsu (1939 b), Mood (1951), Roy (1939) 等によって同時に解かれた．ムードの結果は得られてから 12 年間発表されなかった．最近では James (1954), Olkin (1951), Olkin と Roy (1954) 等が異なった方法で取り扱っている．

基本的な結果は次のように述べられている．

18.6.2　$(x_{1\xi}, \cdots, x_{k\xi},\ \xi = 1, \cdots, n)$ は正規分布 $N(\{\mu_i\};\|\sigma_{ij}\|)$ からの標本で，$\|u_{ij}\|$ はこの標本の内部散乱行列であるとする．固有方程式

$$|u_{ij} - l\sigma_{ij}| = 0 \tag{18.6.20}$$

の根 $0 < l_k < \cdots < l_1$ は次の確率素分で与えられる分布を持つ．

$$\begin{cases} \dfrac{\pi^{\frac{1}{2}k}\left(\prod_{i=1}^{k} l_i\right)^{\frac{1}{2}(n-k-2)} \prod_{i>j=1}^{k} (l_j - l_i) \exp\left(-\frac{1}{2}\sum_{i=1}^{k} l_i\right)}{2^{\frac{1}{2}k(n-1)} \prod_{i=1}^{k} \Gamma\left(\frac{k+1-i}{2}\right)\Gamma\left(\frac{n-i}{2}\right)} dl_1 \cdots dl_k, & 0 < l_k < l_{k-1} < \cdots < l_1 < +\infty \\ 0, & \text{その他.} \end{cases} \tag{18.6.21}$$

18.6.2 は次のようにして示される．$\{u_{ij}\}$ がウィシャート分布 $W(k, n-1, \|\sigma_{ij}\|)$ を持つことはすでにわかっている．さらに正定値 2 次形式の理論［たとえば Birkhoff と MacLane (1953) を参照］より，正則な 1 次変換

$$(x_{i\xi} - \mu_i) = \sum_{j=1}^{k} c_{ij} y_{j\xi}, \qquad i = 1, \cdots, k,\ \ \xi = 1, \cdots, n \tag{18.6.22}$$

を見つけて $(y_{1\xi}, \cdots, y_{k\xi},\ \xi = 1, \cdots, n)$ を $N(\{0\};\|\delta_{ij}\|)$ からの標本とすることができる．ただし $\|\delta_{ij}\|$ は単位行列である．簡単のために行列 $\|c_{ij}\|$, $\|\sigma_{ij}\|$, $\|\delta_{ij}\|$ をそれぞれ c, σ, I, c および σ の逆行列を c^{-1} および σ^{-1}, 行列の転置をダッシュ（$'$）で表わそう．すると

$$c\sigma^{-1}c' = I, \qquad (c^{-1})'\sigma c^{-1} = I. \tag{18.6.23}$$

が成立する．

また $\|u^*_{ij}\|$ を y の標本の内部散乱行列とすれば $\{u^*_{ij}\}$ はウィシャート分布 $W(k, n-1, \|\delta_{ij}\|)$ を持つ．$\|u_{ij}\|$，$\|u^*_{ij}\|$ をそれぞれ，u, u^* で表わせば u と u^* は互いに次のように表わせる．

(18.6.24) $$u = c'u^*c, \qquad u^* = (c^{-1})'uc^{-1}.$$

さて $l_1 > \cdots > l_k$ を (18.6.20)，すなわち

(18.6.25) $$|u - l\sigma| = 0$$

の根としよう．この根のうち 2 つ以上が等しくなる確率は 0 である．行列 $\|u - l\sigma\|$ の左から $(c^{-1})'$，右から c^{-1} を乗じると

(18.6.26) $$(c^{-1})'uc^{-1} - l(c^{-1})'\sigma c^{-1}$$

が得られる．(18.6.23) および (18.6.24) の 2 番目の式を考慮すれば，これは

(18.6.27) $$u^* - lI$$

になる．もちろん (18.6.26) と (18.6.27) の行列式の値は等しい．しかし (18.6.26) の行列式は $|(c^{-1})'|\cdot|u - l\sigma|\cdot|c^{-1}|$ で $|u - l\sigma|$ は $l = l_1, \cdots, l_k$ に対して 0 になる．したがって $|u^* - lI|$ は同じ l の値に対して 0 になる．

このようにして $\{u_{ij}\}$ がウィシャート分布 $W(k, n-1, \|\sigma_{ij}\|)$ を持つとき，$|u - l\sigma| = 0$ の根の分布関数は $|u^* - lI| = 0$ の根の分布関数に一致する．ただし $\{u^*_{ij}\}$ はウィシャート分布 $W(k, n-1, \|\delta_{ij}\|)$ に従う．$\{u^*_{ij}\}$ の確率素分は

(18.6.28) $$\frac{|u^*_{ij}|^{\frac{1}{2}(n-k-2)} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} u^*_{ii}\right)\prod_{i\geqq j=1}^{k} du^*_{ij}}{2^{\frac{1}{2}k(n-1)}\pi^{k(k-1)/4}\Gamma\left(\frac{n-1}{2}\right)\cdots\Gamma\left(\frac{n-k}{2}\right)}$$

である．

正定値 2 次形式の理論より，確率変数の直交行列 $\|e_{ij}\|$ および正の確率変数の組 $f_k < \cdots < f_1$ が存在して

(18.6.29) $$u^* = e'fe$$

となる．ただし e は行列 $\|e_{ij}\|$，f は対角要素 $f_1, \cdots, f_k$ を持つ対角行列である．また e は直交行列だから $e'e = I$ が成立する．よって行列 $\|u^* - lI\|$ において u^*, I のかわりにそれぞれ $e'fe$，$e'e$ で置き換えれば

(18.6.30) $$u^* - lI = e'fe - le'e = e'(f - l)e$$

を得る．(18.6.30) において行列式をとれば，$|e| \neq 0$ だから根 $l_k < \cdots < l_1$ は $f_k < \cdots < f_1$ に等しいことがわかる．ゆえに根 $l_1, \cdots, l_k$ の分布はそれぞれ $f_1, \cdots, f_k$ の分布に

まったく一致する．$f_1, \cdots, f_k$ の分布を見い出すには変換 (18.6.29) を確率素分 (18.6.28) に施して，$f_1, \cdots, f_k$ に関する周辺分布をとればよい．e は直交行列だから $\{e_{ij}\}$ は $\frac{1}{2}k(k-1)$ 個のパラメータの関数になる．これを g_t，$t = 1, \cdots, \frac{1}{2}k(k-1)$ で表わす．変換のヤコビアンは

(18.6.31) $$J = \left|\frac{\partial(u^*_{ij})}{\partial(g_t, f_{i'})}\right| = \overbrace{\left|\begin{array}{c} \dfrac{\partial(u^*_{ij})}{\partial(g_t)} \\ \cdots \\ \dfrac{\partial(u^*_{ij})}{\partial(f_{i'})} \end{array}\right|}^{\frac{1}{2}k(k+1)\text{ 個の列}} \begin{array}{l} \left.\vphantom{\dfrac{a}{b}}\right\} \frac{1}{2}k(k-1)\text{ 個の行} \\ \\ \left.\vphantom{\dfrac{a}{b}}\right\} k\text{ 個の行} \end{array}$$

で，$i \geqq j = 1, \cdots, k$；$i' = 1, \cdots, k$，$t = 1, \cdots, \frac{1}{2}k(k-1)$ となる．さて J の上から $\frac{1}{2}k(k-1)$ 個までのうち t 番目の行の任意の成分は $\sum_{j=1}^{k} q_j f_j$ なる型をしている．ただし q_j は $\{e_{ij}\}$ と e_{ij} の g_t による 1 次微係数だけに依存する．さらに下の k 個の行における任意の要素は $\{e_{ij}\}$ だけに依存している．よって J は $\{e_{ij}\}$ とその $\{g_t\}$ による微係数だけに依存する係数を持つ $f_1, \cdots, f_k$ に関する $\frac{1}{2}k(k-1)$ 次の斉次多項式になる．

任意の $i \neq j$ に対して $f_i = f_j$ ならば変換 (18.6.29) は一意でない．したがって，すべての $i > j$ に対して J の因数として $(f_j - f_i)^{a_{ij}}$ が存在する．ただし a_{ij} は正整数である．ゆえに J は $\prod_{i>j=1}^{k} (f_j - f_i)^{a_{ij}}G$ なる型を持つ．さて a_{ij} は正整数で，$(f_j - f_i)^{a_{ij}}$ なる型の項が $\frac{1}{2}k(k-1)$ 個あり，J は $\frac{1}{2}k(k-1)$ 次の $f_1, \cdots, f_k$ に関する斉次多項式であるから，a_{ij} はすべて 1 で，G は f_i に依存せず，e_{ij} と e_{ij} の g_t による 1 次微係数だけで決まる．このように G は g_t だけに関係するから $G(\{g_t\})$ と書ける．すなわち

(18.6.32) $$J = \prod_{i>j=1}^{k} (f_j - f_i)G(\{g_t\})$$

となる．ここで $|e_{ij}| = 1$ であるから (18.6.29) より

(18.6.33) $$|u^*_{ij}| = |e'|\cdot|f|\cdot|e| = f_1 \cdots f_k$$

を得る．さらに

(18.6.34) $$\sum_{i=1}^{k} u^*_{ii} = \sum_{i=1}^{k} f_i.$$

(18.6.32)，(18.6.33) および (18.6.34) を用いれば，(18.6.28) に変換 (18.6.29) を

適用して次の結果を得る．

(18.6.35) $$KG(g_t)\prod_{t=1}^{\frac{1}{2}k(k-1)} dg_t\left(\prod_{i=1}^{k} f_i\right)^{\frac{1}{2}(n-k-2)}\prod_{i>j=1}^{k}(f_j-f_i)\exp\left(-\frac{1}{2}\sum_{i=1}^{k} f_i\right)\prod_{i=1}^{k} df_i.$$

ただし K は定数．$\{g_t\}$ と $\{f_i\}$ は確率変数の独立な組だから $\{f_i\}$ の p.e. は

(18.6.36) $$C\left\{\left(\prod_{i=1}^{k} f_i\right)^{\frac{1}{2}(n-k-2)}\prod_{i>j=1}^{k}(f_j-f_i)\exp\left(-\frac{1}{2}\sum_{i=1}^{k} f_i\right)\right\}\prod_{i=1}^{k} df_i$$

である．ただし C はあとで決められる基準化定数．{ } 内の関数を $R(k, n, \{f_i\})$ で表わせば次のように書ける．

(18.6.37) $$\frac{1}{C}=\int_{E_k} R(k, n, \{f_i\})\, df_1\cdots df_k.$$

ただし E_k は $0<f_k<f_{k-1}<\cdots<f_1<+\infty$ なる R_k の領域である．(18.6.37) の右辺の積分を $\varphi_k\left[\frac{1}{2}(n-k-2)\right]$ で表わすと，$|u^*_{ij}|^r=\left(\prod_{i=1}^{k} f_i\right)^r$ だから $|u^*_{ij}|$ の r 次のモーメントは次のようになる．

(18.6.38) $$\mathscr{E}(|u^*_{ij}|^r)=\frac{\varphi_k\left(\frac{n-k-2}{2}+r\right)}{\varphi_k\left(\frac{n-k-2}{2}\right)}$$

しかし系 $\{u^*_{ij}\}$ は p.e. (18.6.28) を持つウィシャート分布 $W(k, n-1, \|\delta_{ij}\|)$ に従うから $|u^*_{ij}|$ の r 次のモーメントの値は (18.2.33) において $\|\sigma_{ij}\|$ を $\|\delta_{ij}\|$, n を $n-1$ と置き換えると得られる．ゆえに

(18.6.39) $$\frac{\varphi_k\left(\frac{n-k-2}{2}+r\right)}{\varphi_k\left(\frac{n-k-2}{2}\right)}=2^{rk}\prod_{i=1}^{k}\frac{\Gamma\left(\frac{n-i}{2}+r\right)}{\Gamma\left(\frac{n-i}{2}\right)}.$$

$=-\frac{1}{2}(n-k-2)$ とすれば

(18.6.40) $$C=\frac{1}{\varphi_k\left(\frac{n-k-2}{2}\right)}=\frac{2^{-\frac{1}{2}k(n-k-2)}}{\varphi_k(0)}\prod_{i=1}^{k}\frac{\Gamma\left(\frac{k+2-i}{2}\right)}{\Gamma\left(\frac{n-i}{2}\right)}$$

となる．しかし

(18.6.41) $$\varphi_k(0)=\int_{E_k}\prod_{i>j=1}^{k}(f_j-f_i)\exp\left(-\frac{1}{2}\sum_{i=1}^{k} f_i\right)df_1\cdots df_k$$

である．

変換 $f_k=f'_k,\ f_{k-1}=f'_{k-1}+f'_k, \cdots, f_1=f'_1+f'_k$ を行なえば

(18.6.42) $$\varphi_k(0)=\frac{2}{k}\varphi_{k-1}(1)$$

がわかる．しかし (18.6.40) から $n=k+4$ と置くと

(18.6.43) $$\frac{\varphi_k(0)}{\varphi_k(1)}=2^{-k}\prod_{i=1}^{k}\frac{\Gamma\left(\frac{k+2-i}{2}\right)}{\Gamma\left(\frac{k+4-i}{2}\right)}=\frac{1}{(k+1)!}$$

が得られる．

(18.6.42) および (18.6.43) より

$$\varphi_k(1)=(k+1)!\frac{2}{k}\varphi_{k-1}(1).$$

k を逐次 $k-1, k-2, \cdots, 2$ と置き換えて $\varphi_1(1)=4$ に注意すれば

(18.6.44) $$\varphi_k(1)=2^k(k+1)!\prod_{i=1}^{k}\Gamma(k+1-i)$$

となり，(18.6.43) に代入すると

(18.6.45) $$\varphi_k(0)=2^k\prod_{i=1}^{k}\Gamma(k+1-i)$$

が得られる．しかしガンマ関数に対するルジャンドルの重複公式

(18.6.46) $$\Gamma(k+1-i)=\frac{2^{k-i}}{\sqrt{\pi}}\Gamma\left(\frac{k+1-i}{2}\right)\Gamma\left(\frac{k+2-i}{2}\right)$$

を (18.6.45) に代入し，さらに $\varphi_k(0)$ の値を (18.6.40) に代入すれば

(18.6.47) $$C=\frac{\pi^{\frac{1}{2}k}}{2^{\frac{1}{2}k(n-1)}\prod_{i=1}^{k}\Gamma\left(\frac{k+1-i}{2}\right)\Gamma\left(\frac{n-i}{2}\right)}$$

が得られる．この値を (18.6.36) に代入すれば最終的に $f_1, \cdots, f_k$ の分布が得られる．しかし $f_1, \cdots, f_k$ はそれぞれ確率変数 $l_1, \cdots, l_k$ に等しい．ゆえに $l_1, \cdots, l_k$ の分布は単に (18.6.36) において $f_1, \cdots, f_k$ を $l_1, \cdots, l_k$ に置き換えることによって与えられる．これは (18.6.21) である．これで **18.6.2** の議論を終える．

(18.6.21) で与えられた $l_1, \cdots, l_k$ の分布は空の場合に対する分布と呼ばれることがある．一般の場合は標本が抽出された正規分布の共分散行列が (18.6.20) の行列 $\|\sigma_{ij}\|$ と異なっている場合である．このようなより一般の場合における固有値の分布は James

(1960) により見い出されている．それは (18.6.21) よりはるかに複雑になっている．

18.7 判別分析

(a) 2標本の場合

本節では次のような問題を考えよう．k 次元分布からの 2 つの標本がある．幾何学的には，これらは k 次元ユークリッド空間における 2 つの標本クラスターとして表現される．この 2 つの標本を直線上に直交射影し，射影された標本の標本内偏差に対する標本間偏差を最大にしたい．問題はこの値を最大にするような射影の方向を求めることである．いい換えれば，2 つの標本クラスターを 1 次元空間に引き戻して射影されたあとの 2 標本クラスターが標本内偏差に対してできるだけ離れるようにすることである．実際的な立場からすれば次のようになる．2 標本が，k 個の成分のうち 1 つあるいはいくつかの少ない成分からなる空間へ射影した場合は，たとえ分離されていなくとも，適度に分離されるような 1 次元空間への射影方向が見い出されたならば，k 個の成分を持つベクトルの適当な線形変換によって 2 つの分布からの標本を判別できたといえよう．この 2 つの標本はそのベクトルの k 個の成分を観測したときに判別されたことになる．この問題は Fisher (1938) が最初に考えた．この問題の解から発展した統計的な分析方法を判別分析という．さてこの問題に対して厳密な考察を傾けよう．

$(x_{1\xi_\gamma}^{(\gamma)}, \cdots, x_{k\xi_\gamma}^{(\gamma)})$, $\xi_\gamma = 1, \cdots, n_\gamma$), $\gamma = 1, 2$ を 2 つの標本とする．ただし $n_1 > k$, $n_2 > k$. 2 標本の平均値ベクトル，内部散乱行列をそれぞれ $(\bar{x}_1^{(\gamma)}, \cdots, \bar{x}_k^{(\gamma)})$, $\gamma = 1, 2$, $\|u_{ij}^{(\gamma)}\|$, $\gamma = 1, 2$ で表わし，$\|u_{ij}^{(\gamma)}\|$, $\gamma = 1, 2$ は確率 1 で正則であるとする．2 標本を合わせた標本の標本平均，内部散乱行列をそれぞれ $(\bar{x}_1, \cdots, \bar{x}_k)$, $\|u_{ij}\|$ とし，$\|u_{ij}^W\| = \|u_{ij}^{(1)} + u_{ij}^{(2)}\|$ を 2 標本に対する標本内散乱行列とする．行列 $\|u_{ij}^B\| = \|u_{ij} - u_{ij}^W\|$ は標本間散乱行列である．任意のベクトル $(c_1, \cdots, c_k)$ に対して

(18.7.1)
$$z_{\xi_\gamma}^{(\gamma)} = \sum_{i=1}^{k} c_i x_{i\xi_\gamma}^{(\gamma)}, \qquad \xi_\gamma = 1, \cdots, n_\gamma, \quad \gamma = 1, 2$$

とすると $(z_1^{(1)}, \cdots, z_{n_1}^{(1)})$ と $(z_1^{(2)}, \cdots, z_{n_2}^{(2)})$ はそれぞれ大きさを無視すれば，もとの k 次元標本を，もとの k 次元空間における方向余弦が $(c_1, \cdots, c_k)$ に比例するように直線上に射影して得られる 1 次元標本となる．$\bar{z}^{(1)}$, $\bar{z}^{(2)}$ を z の 2 つの標本平均，$\bar{z}$ を合わせた標本の標本平均として次を定義する．

(18.7.2)
$$S_W = \sum_{\gamma=1}^{2} \sum_{\xi_\gamma=1}^{n_\gamma} (z_{\xi_\gamma}^{(\gamma)} - \bar{z}^{(\gamma)})^2, \qquad S_B = \sum_{\gamma=1}^{2} n_\gamma (\bar{z}^{(\gamma)} - \bar{z})^2.$$

S を z の 2 つの標本を合わせた標本の散乱とすれば $S = S_W + S_B$ が成立する．S_W は標本内散乱成分，S_B は標本間散乱成分である．さて問題は S_W の固定された値に対して S_B を最大（すなわち $S_W/(S_W + S_B)$ を最小）にするような $(c_1, \cdots, c_k)$ を定めることである．

この問題に関する基本的な結果は次のように述べられている．

18.7.1　$\|u_{ij}^W\|$, $\|u_{ij}^B\|$ はともに標本の大きさが k を越える k 次元分布からの 2 つの標本の標本内，標本間散乱行列とし，$\|u_{ij}^W\|$ は確率 1 で正定値とする．S_B, S_W を (18.7.2) で定義すれば S_W が固定された値 $C \neq 0$ を持つとき，比

(18.7.3)
$$Q = \frac{S_W}{S_W + S_B}$$

を最小にする $(c_1, \cdots, c_k)$ の値（たとえばこれを $(c_1', \cdots, c_k')$ とする）は方程式

(18.7.4)
$$\sum_{j=1}^{k} (u_{ij}^B - l_1 u_{ij}^W) c_j = 0, \qquad i = 1, \cdots, k$$

の解である．ただし l_1 は

(18.7.5)
$$l_1 = \frac{n_1 n_2}{n_1 + n_2} \sum_{i,j=1}^{k} u_W^{ij} (\bar{x}_i^{(1)} - \bar{x}_i^{(2)})(\bar{x}_j^{(1)} - \bar{x}_j^{(2)}) = m^2$$

で与えられる固有方程式

(18.7.6)
$$|u_{ij}^B - l u_{ij}^W| = 0$$

の 0 でない根である．ここに

(18.7.7)
$$m = \frac{1}{\sqrt{C}} \sum_{i=1}^{k} c_i' b_i, \qquad \|u_W^{ij}\| = \|u_{ij}^W\|^{-1}$$

で，b_i は (18.4.14) で与えられている．さらに $S_W = C \neq 0$ に対する Q の最小値は $1/(1 + l_1)$ である．

18.7.1 を証明するには，まず次に注意する．

(18.7.8)
$$S_B = \sum_{\gamma=1}^{2} n_\gamma \left[\sum_{i=1}^{k} c_i(\bar{x}_i^{(\gamma)} - \bar{x}_i)\right]^2 = \sum_{i,j=1}^{k} u_{ij}^B c_i c_j$$
$$S_W = \sum_{\gamma=1}^{2} \sum_{\xi_\gamma=1}^{n} \left[\sum_{i=1}^{k} c_i(x_{i\xi_\gamma}^{(\gamma)} - \bar{x}_i^{(\gamma)})\right]^2 = \sum_{i,j=1}^{k} u_{ij}^W c_i c_j.$$

条件 $S_W = C \fallingdotseq 0$ の下で $S_W/(S_W + S_B)$ を最小にすることは（ラグランジュ乗数 l を用いて）$S_B + (C - S_W)l$ を最大にすることと同値である．$S_B + (C - S_W)l$ を $(c_1, \cdots, c_k)$ と l の関数として φ で表わせば，φ の最大値は方程式

(18.7.9)
$$\frac{\partial\varphi}{\partial c_i} = 0, \qquad i = 1, \cdots, k$$
$$\frac{\partial\varphi}{\partial l} = 0$$

の解で与えられる．この方程式は

(18.7.10)
$$\sum_{j=1}^{k}(u_{ij}^B - l u_{ij}^W)c_j = 0, \qquad i = 1, \cdots, k$$

と書かれる．(18.7.10) に対する $(0, \cdots, 0)$ 以外の解 $(c_1', \cdots, c_k')$ を得るためには

(18.7.11)
$$|u_{ij}^B - l u_{ij}^W| = 0$$

を解く必要がある．

18.4(b) 節より
$$u_{ij}^B = b_i b_j$$
となる．ただし (18.4.14) で見たように

(18.7.12)
$$b_i = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}(\bar{x}_i^{(1)} - \bar{x}_i^{(2)}).$$

よって (18.7.11) は
$$l^k \left| u_{ij}^W - \frac{b_i b_j}{l} \right| = 0$$
と同値．これは $\|u_W^{ij}\| = \|u_{ij}^W\|^{-1}$ として

(18.7.13)
$$l^{k-1}|u_{ij}^W| \cdot \left[l - \sum_{i,j=1}^{k} u_W^{ij} b_i b_j\right] = 0$$

となる．したがって (18.7.11) の 0 でない根は

(18.7.14)
$$l_1 = \sum_{i,j=1}^{k} u_W^{ij} b_i b_j$$

で与えられる．ゆえに (18.7.10) の解 $(c_1', \cdots, c_k')$ は

(18.7.15)
$$\sum_{j=1}^{k}(u_{ij}^B - l_1 u_{ij}^W)c_j' = 0, \qquad i = 1, \cdots, k$$

を満足する．
$$S_W = \sum_{i,j=1}^{k} u_{ij}^W c_i' c_j' = C$$
に注意して，(18.7.15) に c_i' を乗じて i に関して加えれば

(18.7.16)
$$\sum_{i,j=1}^{k} u_{ij}^B c_i' c_j' - l_1 C = 0$$

が得られる．よって (18.7.6) で述べられたように (18.7.12) の b_i を用いて
$$l_1 = \frac{1}{C}\left(\sum_{i=1}^{k} c_i' b_i\right)^2$$
となる．

(18.7.8) の c_i のかわりに c_i' を代入して，S_B，S_W を求め，この値を Q に代入する．ここで (18.7.16) を利用すれば，Q の最小値は $1/(1 + l_1)$ になる．これで **18.7.1** の議論を終える．

次元が大きい正規分布からの 2 つの標本の場合，すなわち，n_1，n_2 が k よりも小さい場合に対する判別分析の問題は Dempster (1958) によって取り扱われている．

(b) いくつかの標本の場合

2 標本問題においては，2 つの k 次元標本を直線上に射影して得られる 1 次元標本に関して標本内散乱に対する標本間散乱をできるだけ大きくする射影（の方向）を示した．3 つの k 次元標本の場合には，これらの標本を 2 次元空間（普通の平面）に射影して，2 次元標本を合わせた標本の散乱を 3 つの射影された標本の 標本内散乱に比べて できるだけ大きくしたい．一般には $(s+1)$ 個の k 次元標本 $(s + 1 \leqq k)$ があったとき，この標本を s 次元ユークリッド空間上へ射影し，その結果できる標本に対して，$(s+1)$ 個の標本を合わせた標本の散乱が，$(s+1)$ 個の s 次元標本の標本内散乱に比較して，できるだけ大きくなるようにしたい．

もっと明確に述べよう．$(x^{(\gamma)}_{1\xi_\gamma}, \cdots, x^{(\gamma)}_{k\xi_\gamma}; \xi_\gamma = 1, \cdots, n_\gamma)$, $\gamma = 1, \cdots, s+1$ を $s+1$ 個の k 次元標本で，$s+1 \leqslant k$ とする．$\|u^{(\gamma)}_{ij}\|$, $\gamma = 1, \cdots, s+1$ はそれぞれ標本平均 $(\bar{x}^{(\gamma)}_1, \cdots, \bar{x}^{(\gamma)}_k)$ のまわりの内部散乱行列で，$\|u_{ij}\|$ は合わせた標本の内部散乱行列とする．$\|u^W_{ij}\| = \|u^{(1)}_{ij} + \cdots + u^{(s+1)}_{ij}\|$ は標本内散乱で，確率 1 で正則とする．$\|u^B_{ij}\| = \|u_{ij} - u^W_{ij}\|$ は標本間散乱行列である．$\|u^W_{ij}\|$ が確率 1 で正則ならば $\|u_{ij}\|$ もそうである．$z^{(\gamma)}_{\xi_\gamma}$, $\gamma = 1, \cdots, s+1$ を (18.7.1) で定義する．ただし $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ は 1 次独立である．γ $(\gamma = 1, \cdots, s+1)$ 番目の標本に対して，射影された s 次元標本の座標を

(18.7.17) $$z^{(\gamma)}_{p\xi} = \sum_{i=1}^{k} c_{ip} x^{(\gamma)}_{i\xi}, \qquad p = 1, \cdots, s, \quad \xi = 1, \cdots, n_\gamma$$

とする．$\|\bar{u}^{(\gamma)}_{pq}\|$ $(\gamma = 1, \cdots, s+1)$ を z の γ 番目の標本の内部散乱行列，$\|\bar{u}_{pq}\|$ を z の合わせた標本の内部散乱行列とする．z の $(s+1)$ 個の標本の標本内散乱行列 $\|\bar{u}^{(1)}_{pq} + \cdots + \bar{u}^{(s+1)}_{pq}\|$ を $\|\bar{u}^W_{pq}\|$ で，z の標本間散乱行列 $\|\bar{u}_{pq} - \bar{u}^W_{pq}\|$ を $\|\bar{u}^B_{pq}\|$ で表わす．すると問題は散乱 $|\bar{u}^W_{pq}|$ に対して散乱 $|\bar{u}_{pq}|$ の値を最大にするような s 個の 1 次独立なベクトル $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ を求めることである．これは次と同値である．すなわち $(s+1)$ 個の標本を s 次元空間 R_s 上へ射影して，R_s において合わせた標本の散乱に対する標本内散乱の比がもとの空間 R_k での合わせた標本の散乱に対する標本内散乱の比に等しくなるように射影の方向を定めることである．すなわち

(18.7.18) $$\frac{|\bar{u}^W_{pq}|}{|\bar{u}_{pq}|} = \frac{|u^W_{ij}|}{|u_{ij}|}$$

となるベクトル $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ を見つけたい．

このベクトルを求めるには 2 標本の場合と同様に任意のベクトル $(c_1, \cdots, c_k)$ に対して，S_B, S_W を (18.7.8) のように定義することから始めよう．そうすれば S_W が固定された値 $C \neq 0$ を持つという条件の下で，S_B が定留関数となるようなベクトルの値が見つかる．この値は方程式 (18.7.9) で与えられる．ただし $\varphi = S_B + (C - S_W)l$. この φ に対する方程式は (18.7.10) で与えられる．もちろん $\|u^B_{ij}\|$, $\|u^W_{ij}\|$ は 2 標本の場合でなく $(s+1)$ 個の標本に対する標本間，標本内散乱である．$(s+1)$ 個の標本の場合に (18.7.10) が解けるための条件は (18.7.11) を $(s+1)$ 個の標本の場合に変形すればよい．すなわち

(18.7.19) $$|u^B_{ij} - l u^W_{ij}| = 0.$$

である．さて $n_1 + \cdots + n_{s+1} - (s+1) \geqslant k$ で $s+1 \leqslant k$ ならば，$\|u^B_{ij}\|$, $\|u_{ij}\|$ の階数はそれぞれ s, k（確率 1 で）であることが確かめられる．よって $\|u^W_{ij}\|$ は正定値，$\|u^B_{ij}\|$ は非負定値になる．この 2 つの行列を u^W, u^B で表わせば，行列論 [たとえば Birkhoff と McLane (1953) を参照] により，実数要素を持つ正則行列 $\|e_{ij}\| = e$ および確率 1 ですべて異なる s 個の実数，すなわち $0 < l_s < \cdots < l_1 < +\infty$ が存在して

(18.7.20) $$e' u^W e = I, \qquad e' u^B e = L$$

となる．ただし L は $l_1, \cdots, l_s, 0, \cdots, 0$ を対角要素に持つ $k \times k$ 次元の対角行列である．したがって行列 $(u^B - l u^W)$ の左から e', 右から e を乗じれば

(18.7.21) $$(e' u^B e - l e' u^W e)$$

を得る．この行列の行列式をとれば

(18.7.22) $$|e' u^B e - l e' u^W e| = |L - lI| = l^{k-s}(l_1 - l) \cdots (l_s - l).$$

しかし

(18.7.23) $$|e' u^B e - l e u^W e| = |e'| \cdot |u^B - l u^W| \cdot |e|$$

で $|e| = |e'| \neq 0$ だから $|u^B - l u^W|$ が 0 になる l の値は $|e' u^B e - l e' u^W e|$ を 0 にする l の値に一致する．しかし (18.7.22) より，$|e' u^B e - l e' u^W e|$ の 0 でない根（固有値）は $l_1, \cdots, l_s$ である．さて $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ を (18.7.10) の $(s+1)$ 個の標本の場合の解（固有ベクトル）とする．すなわち $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ は

(18.7.24) $$\sum_{j=1}^{k} (u^B_{ij} - l_p u^W_{ij}) c_{jp} = 0, \qquad i = 1, \cdots, k$$

(18.7.24 a) $$\sum_{i,j=1}^{k} u^W_{ij} c_{ip} c_{jp} = C$$

を満たす．方程式 (18.7.24 a) は $C - S_W = 0$ にすぎない．ただし S_W は $(c_1, \cdots, c_k) = (c_{1p}, \cdots, c_{kp})$ のときの値である．

$(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ が求めるベクトルであることを示そう．方程式 (18.7.24) に c_{ip} を乗じて i に関して加えると (18.7.24 a) より

(18.7.25) $$\sum_{i,j=1}^{k} u^B_{ij} c_{ip} c_{jp} = l_p C$$

を得る．さて (18.7.24) に c_{iq}, $q \neq p$ を乗じて i に関して加えると

(18.7.26) $$\sum_{i,j=1}^{k} u^B_{ij} c_{jp} c_{iq} - l_p \sum_{i,j=1}^{k} u^W_{ij} c_{jp} c_{iq} = 0.$$

同様に

(18.7.27)
$$\sum_{i,j=1}^{k} u_{ij}^{B} c_{jq} c_{ip} - l_q \sum_{i,j=1}^{k} u_{ij}^{W} c_{jq} c_{ip} = 0.$$

$l_p \neq l_q$ だから，(18.7.26) と (18.7.27) の差をとると

(18.7.28)
$$\sum_{i,j=1}^{k} u_{ij}^{W} c_{jp} c_{iq} = 0.$$

よって (18.7.26) または (18.7.27) より

(18.7.29)
$$\sum_{i,j=1}^{k} u_{ij}^{B} c_{jp} c_{iq} = 0.$$

さて次が確かめられる．

(18.7.30)
$$\bar{u}_{pq}^{W} = \sum_{i,j=1}^{k} u_{ij}^{W} c_{ip} c_{jq} = \delta_{pq} C$$

(18.7.31)
$$\bar{u}_{pq} = \bar{u}_{pq}^{W} + \bar{u}_{pq}^{B} = \sum_{i,j=1}^{k} (u_{ij}^{W} + u_{ij}^{B}) c_{ip} c_{jq} = \delta_{pq} C(1 + l_p)$$

ただし $\delta_{pq} = \begin{cases} 1 & (p = q \text{ のとき}) \\ 0 & (p \neq q \text{ のとき}) \end{cases}$. ゆえに

(18.7.32)
$$\frac{|\bar{u}_{pq}^{W}|}{|\bar{u}_{pq}|} = \frac{1}{(1 + l_1)\cdots(1 + l_s)}$$

が得られる．

さてもとの k 次元空間における散乱の比 $|u_{ij}^{W}|/|u_{ij}|$ を考えよう．(18.7.20) より

$$|e' u^{W} e| = |e'| \cdot |u^{W}| \cdot |e| = 1$$

だから $|u_{ij}^{W}| = 1/|e_{ij}|^2$ となる．$|u_{ij}| = |u_{ij}^{W} + u_{ij}^{B}|$ より (18.7.20) から同様に $|u_{ij}| = (1 + l_1)\cdots(1 + l_s)/|e_{ij}|^2$. ゆえに

(18.7.33)
$$\frac{|u_{ij}^{W}|}{|u_{ij}|} = \frac{1}{(1 + l_1)\cdots(1 + l_s)}$$

となり (18.7.32) および (18.7.33) の左辺における散乱の比に等しくなる．この共通な値は $[1/(1 + l_1)\cdots(1 + l_s)]$ である．

要約すると次のようになる．

18.7.2　$(x_{1\xi_\gamma}^{(\gamma)}, \cdots, x_{k\xi_\gamma}^{(\gamma)}\,;\, \xi_\gamma = 1, \cdots, n_\gamma)$, $\gamma = 1, \cdots, s + 1$, $n_1 + \cdots + n_{s+1} - (s + 1) \geqslant k$, $s + 1 \leqslant k$ は $(s + 1)$ 個の独立な k 次元標本で，合わせた標本の内部散乱に対する標本内散乱の比を $|u_{ij}^{W}|/|u_{ij}|$ とする．ただし確率 1 で $|u_{ij}^{W}| \neq 0$. この標本点を (18.7.17) によって s 次元空間上へ線形変換して得られる標本に対する合わせた標本の内部散乱と標本内散乱との比を $|\bar{u}_{pq}^{W}|/|\bar{u}_{pq}|$ とする．(18.7.24 a) を満たし，2 つの比 $|u_{ij}^{W}|/|u_{ij}|$ と $|\bar{u}_{pq}^{W}|/|u_{pq}|$ が等しくなるような固有ベクトル $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s < k$ は (18.7.24) の解である．ただし固有値 $0 < l_s < \cdots < l_1 < +\infty$ は (18.7.19) の 0 でない根である．この比の共通な値は $1/[(1 + l_1)\cdots(1 + l_s)]$ である．固有値 l_p, $p = 1, \cdots, s$ は (18.7.25) で与えられ，確率 1 ですべて異なる．$p \neq q$ となる任意の 2 つの固有ベクトル $(c_{1p}, \cdots, c_{kp})$ と $(c_{1q}, \cdots, c_{kq})$ はともに条件 (18.7.28), (18.7.29) を満たす．

この定理の核心は次の事柄を意味している．s 個の固有ベクトル $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ を用いて，k 次元標本点を s 次元空間へ線形射影すれば R_s において射影された散乱の比 $|\bar{u}_{pq}^{W}|/|\bar{u}_{pq}|$ はもとの k 次元空間 R_k における散乱の比 $|u_{ij}^{W}|/|u_{ij}|$ に等しくなる．これは次と同値である．すなわち k 次元標本点を s 次元ユークリッド空間に射影して，標本間散乱 $|\bar{u}_{pq}^{W}|$ に対して，合わせた標本の散乱 $|\bar{u}_{pq}|$ を，できるだけ大きくなるようにすることである．$s + 1$ 個の標本を 1 次元の点へ射影して合わせた標本の散乱が標本内散乱に対して最大になるようにしようとすれば (18.7.17) では最大の固有値 l_1 に対する固有ベクトル $(c_{11}, \cdots, c_{k1})$ を用いればよい．1 次元直線上へ射影した合わせた標本散乱に対する標本内散乱の比は $1/(1 + l_1)$ である．同様に標本を t $(< s)$ 次元空間へ射影して比を最大にしようとすれば，固有値 $l_1, \cdots, l_t$ に対応する固有ベクトル $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, t$ を用いればよい．このとき t 次元空間での合わせた標本の散乱に対する標本内散乱の比は $1/[(1 + l_1)\cdots(1 + l_t)]$ になる．

次の事柄も重要である．いま $s + 1$ 個の標本がいずれも同じ共分散行列 $\|\sigma_{ij}\|$ を持つ k 次元正規分布から抽出されたものであると仮定しよう．このときこれらの正規分布もまた同一の平均値ベクトルを持つという仮説を検定したい．散乱の比 $|u_{ij}^{W}|/|u_{ij}|$ はこの仮説検定に対するネイマン＝ピヤソンの尤度比と同値になる．さらに仮説が正しければ，この比は (18.5.3) で，$m = n_1 + \cdots + n_{s+1} - (s + 1)$ とした R_s と同じ分布を持つ．

18.8　判別分析における固有値の分布

18.7 節の判別分析問題においては固有値 $l_s < \cdots < l_1$ およびこれに対応する固有ベクトル $(c_{1p}, \cdots, c_{kp})$, $p = 1, \cdots, s$ が基本的な役割を果たしていた．k 次元標本クラスターが k 次元空間において離れている度合は固有値の絶対値（の大きさ）に依存する．すなわち固有値が大きければ大きいほど標本クラスターが離れている度合も大きい．標本が同一の k 次元正規分布からの確率標本ならば固有値は分布を持つ．この分布は容易に求められる．特に k 次元正規分布からの 2 標本の場合，標本内散乱の全散乱に対する比は値 $1/(1+l_1)$ を持つ．ただし l_1 は唯一の固有値である．しかも 18.4 (b) 節より $1/(1+l_1)$ は (18.4.15) で定義された $R_1^{\prime}$ に一致する．よって $R_1^{\prime}$ と同じ分布を持つ．さらに 2 標本問題に対しては $(n_1+n_2-2)/l_1$ がホテリングの T^2 になることも注意しておこう．

この節では $(s+1)$ 個の標本がすべて同一の正規分布から抽出されたときの固有値 $l_1, \cdots, l_s$ の標本分布を考察しよう．2 つの重要な場合が考えられる．（i）$s+1>k$ のとき，すなわち固有値の個数が k の場合と（ii）$s+1 \leqslant k$, すなわち固有値の個数が k 以下の s の場合である．

（a） k 個の固有値の場合

これらの固有値について基本的な分布論を導くに際して，まず次の一般的な結果を示しておくと便利である．

18.8.1　$\{v_{ij}\}$, $\{v_{ij}^{\prime}\}$ はそれぞれウィシャート分布 $W(k, n, \|\sigma_{ij}\|)$, $W(k, n', \|\sigma_{ij}\|)$ を持つ確率変数の独立な組とする．ただし $n, n' > k$. このとき方程式

(18.8.1)
$$|v_{ij}^{\prime} - (v_{ij} + v_{ij}^{\prime})g| = 0$$

の根 $0 < g_k < \cdots < g_1 < +\infty$ は次の確率素分で与えられる分布を持つ．

(18.8.2)
$$\begin{cases} K\left[\prod_{i=1}^{k}(1-g_i)\right]^{\frac{1}{2}(n-k-1)}\left[\prod_{i=1}^{k} g_i\right]^{\frac{1}{2}(n'-k-1)} \prod_{i>j=1}^{k}(g_j - g_i)\,dg_1\cdots dg_k, \\ \qquad\qquad 0 < g_k < g_{k-1} < \cdots < g_1 < 1 \text{ なる領域} \\ 0, \qquad\qquad \text{その他.} \end{cases}$$

ただし

$$K = \pi^{\frac{1}{2}k} \prod_{i=1}^{k} \frac{\Gamma\left(\frac{n+n'+1-i}{2}\right)}{\Gamma\left(\frac{n+1-i}{2}\right)\Gamma\left(\frac{n'+1-i}{2}\right)\Gamma\left(\frac{k+1-i}{2}\right)}.$$

18.6.2 と同様な方法を用いれば，**18.8.1** の証明は

(18.8.1 *a*)
$$|v_{ij}^{\prime *} - (v_{ij}^{*} + v_{ij}^{\prime *})g| = 0$$

の根が (18.8.2) の確率素分を持つことを示すことと同値になる．ただし $\{v_{ij}^{*}\}$, $\{v_{ij}^{\prime *}\}$ はウィシャート分布 $W(k, n, \|\delta_{ij}\|)$, $W(k, n', \|\delta_{ij}\|)$ を持つ確率変数の独立な組である．よって次のように議論しよう．

$\|v_{ij}^{*} + v_{ij}^{\prime *}\|$ と $\|v_{ij}^{\prime *}\|$ はともに正定値（確率 1 で）だから，実数値を要素に持つ正則行列 $\|e_{ij}\| = e$ と対角要素が確率 1 ですべて異なっている対角行列 m, すなわち $1 > m_1 > \cdots > m_k > 0$ が存在して，次を満たす（$\|v_{ij}^{*} + v_{ij}^{\prime *}\|$ を $v^* + v'^*$, $\|v_{ij}^{\prime *}\|$ を v'^* で表わす）．

(18.8.3)
$$v^* + v'^* = e'e, \qquad v'^* = e'me.$$

しかし，(18.8.3) は次と同値である．

(18.8.3 *a*)
$$v'^* = e'me, \qquad v^* = e'(I-m)e.$$

(18.8.3) から方程式 (18.8.1 *a*) は

(18.8.4)
$$|e'me - ge'e| = 0$$

すなわち

(18.8.4 *a*)
$$|e'|\cdot|m - gI|\cdot|e| = 0$$

と書ける．ゆえに根 $g_1, \cdots, g_k$ はそれぞれ $m_1, \cdots, m_k$ に等しい．

さて $\{v_{ij}^{*}\}$ と $\{v_{ij}^{\prime *}\}$ との確率素分は

(18.8.5)
$$A|v_{ij}^{*}|^{\frac{1}{2}(n-k-1)}|v_{ij}^{\prime *}|^{\frac{1}{2}(n'-k-1)} e^{-\frac{1}{2}\sum_{i=1}^{k}(v_{ii}^{*}+v_{ii}^{\prime *})} \prod_{i \geqslant j=1}^{k} dv_{ij}^{*}\,dv_{ij}^{\prime *}$$

になる．ただし A は基準化定数．

(18.8.5) に変換 (18.8.3 *a*) を適用すればヤコビアン J の型は

(18.8.6)
$$J = \begin{array}{c} \overbrace{\qquad\quad}^{k^2\text{ 列}}\ \overbrace{\qquad\quad}^{k\text{ 列}} \\ \left|\begin{array}{c:c} \dfrac{\partial(v_{ij}^{*})}{\partial(e_{i'j'})} & \dfrac{\partial(v_{ij}^{*})}{\partial(m_{i'})} \\ \hdashline \dfrac{\partial(v_{ij}^{\prime *})}{\partial(e_{i'j'})} & \dfrac{\partial(v_{ij}^{\prime *})}{\partial(m_{i'})} \end{array}\right| \end{array} \begin{array}{l} \\ \left.\right\}\ \frac{1}{2}k(k+1)\text{ 行} \\ \left.\right\}\ \frac{1}{2}k(k+1)\text{ 行} \end{array}$$

になる．上の $\frac{1}{2}k(k+1)$ 個の行に，それぞれ下の $\frac{1}{2}k(k+1)$ 個の行を加えれば，J の右上半ブロックの要素は $\partial(v_{ij}^{*} + v_{ij}^{\prime *})/\partial(m_{i'})$ なる型となる．これは (18.8.3) よりすべて 0 になる．m の項を含む J の要素は左下半ブロックだけで，この要素はいずれも e_{ij}

だけに依存する係数を持つ $m_1, \cdots, m_k$ の線形斉次型をしている．よってラプラスの方法［Bôcher (1907) を参照］を用いて J を上の $\frac{1}{2}k(k+1)$ 個の行に関して展開すれば，下の $\frac{1}{2}k(k+1)$ 個の行から構成されるすべての小行列は右下半ブロックから選ばれる $\frac{1}{2}k(k-1)$ 個の列を持つ．したがってラプラス表現における各項は $\frac{1}{2}k(k-1)$ 次の $m_1, \cdots, m_k$ に関する斉次多項式になる．$m_i = m_j$, $i > j$ ならば変換 (18.8.3 a) が定まらず，$J = 0$．よって $(m_j - m_i)^{a_{ij}}$ は J の因数である．ただし a_{ij} は正整数．ゆえに J は $G \prod_{i>j=1}^{k} (m_j - m_i)^{a_{ij}}$ なる型になる．ここに G は e_{ij} だけに関係する．しかし J は $\frac{1}{2}k(k-1)$ 次の多項式だから各 a_{ij} は $a_{ij} = 1$ になる．よって

(18.8.7)
$$J = G \prod_{i>j=1}^{k} (m_j - m_i).$$

(18.8.3) および (18.8.3 a) より次がわかる．

(18.8.8)
$$|v'^{*}_{ij}| = |e_{ij}|^2 \prod_{i=1}^{k} m_i, \qquad |v^{*}_{ij}| = |e_{ij}|^2 \prod_{i=1}^{k} (1 - m_i)$$
$$\sum_{i=1}^{k} (v^{*}_{ii} + v'^{*}_{ij}) = k.$$

よって (18.8.5) に (18.8.7) および (18.8.8) の結果を適用すれば $\{e_{ij}\}$ と $\{m_i\}$ は確率変数の独立な組で，しかも $\{m_i\}$ の確率素分は $0 < m_k < \cdots < m_1 < 1$ なる R_k の領域 E_k 上で

(18.8.9)
$$K\left\{\left[\prod_{i=1}^{k} (1 - m_i)\right]^{\frac{1}{2}(n-k-1)} \left[\prod_{i=1}^{k} m_i\right]^{\frac{1}{2}(n'-k-1)} \prod_{i>j=1}^{k} (m_j - m_i)\right\} dm_1 \cdots dm_k,$$

その他で 0 になる．ただし K はあとで定まる定数．

｛　｝内の E_k 上の積分を $\varphi_k\left(\frac{1}{2}(n-k-1),\ \frac{1}{2}(n'-k-1)\right)$ で表わせば

(18.8.10)
$$\frac{1}{K} = \varphi_k\left(\frac{n-k-1}{2},\ \frac{n'-k-1}{2}\right)$$

を得る．$\{v^{*}_{ij}\}$, $\{v'^{*}_{ij}\}$ は独立なウィシャート分布 $W(k, n, \|\delta_{ij}\|)$, $W(k, n', \|\delta_{ij}\|)$ を持つから (18.2.33) より

(18.8.11)
$$\mathscr{E}(|v^{*}_{ij}|^{r} |v'^{*}_{ij}|^{-r}) = \prod_{i=1}^{k} \frac{\Gamma\left(\frac{n+1-i}{2} + r\right) \Gamma\left(\frac{n'+1-i}{2} - r\right)}{\Gamma\left(\frac{n+1-i}{2}\right) \Gamma\left(\frac{n'+1-i}{2}\right)}$$

と書ける．しかし (18.8.8) より次のようになる．

(18.8.12)
$$\mathscr{E}(|v^{*}_{ij}|^{r} |v'^{*}_{ij}|^{-r}) = \mathscr{E}\left(\left[\prod_{i=1}^{k} (1 - m_i)\right]^{r} \left[\prod_{i=1}^{k} m_i\right]^{-r}\right)$$
$$= \frac{\varphi_k\left(\frac{n-k-1}{2} + r, \frac{n'-k-1}{2} - r\right)}{\varphi_k\left(\frac{n-k-1}{2}, \frac{n'-k-1}{2}\right)}.$$

(18.8.11) および (18.8.12) において $r = -\frac{1}{2}(n-k-1)$ とすれば

(18.8.13)
$$\frac{\varphi_k\left(0, \frac{n+n'-2k-2}{2}\right)}{\varphi_k\left(\frac{n-k-1}{2}, \frac{n'-k-1}{2}\right)} = \prod_{i=1}^{k} \frac{\Gamma\left(\frac{k+2-i}{2}\right) \Gamma\left(\frac{n+n'-k-i}{2}\right)}{\Gamma\left(\frac{n+1-i}{2}\right) \Gamma\left(\frac{n'+1-i}{2}\right)}$$

を得る．いまや K を求める問題は $M = \frac{1}{2}(n + n' - 2k - 2)$ としたときの $\varphi_k(0, M)$ を求めることに帰着した．次式に注意しよう．

(18.8.14)
$$\varphi_k(0, M) = \int_{E_k} \left(\prod_{i=1}^{k} m_i\right)^{M} \prod_{i>j=1}^{k} (m_j - m_i)\, dm_1 \cdots dm_k.$$

$m_1 = t$, $m_2 = tm_2'$, $\cdots$, $m_k = tm_k'$ とすれば

(18.8.15)
$$\varphi_k(0, M) = \frac{1}{k\left(M + \frac{k+1}{2}\right)} \varphi_{k-1}(1, M).$$

(18.8.13) より

(18.8.16)
$$\frac{\varphi_k(0, M)}{\varphi_k(1, M-1)} = \prod_{i=1}^{k} \frac{\left(M + \frac{k-i}{2}\right)}{\left(\frac{k+2-i}{2}\right)}.$$

(18.8.15) および (18.8.16) で定義された 2 パラメータの差分方程式系を解くと

(18.8.17)
$$\varphi_k\left(0, \frac{n+n'-2k-2}{2}\right)$$
$$= \pi^{-\frac{1}{2}k} \prod_{i=1}^{k} \frac{\Gamma\left(\frac{k+2-i}{2}\right) \Gamma\left(\frac{n+n'-k-i}{2}\right) \Gamma\left(\frac{k+1-i}{2}\right)}{\Gamma\left(\frac{n+n'+1-i}{2}\right)}.$$

これを (18.8.13) に代入して，$1/\varphi_k\left(\frac{1}{2}(n-k-1), \frac{1}{2}(n'-k-1)\right)$ の値を求めれば (18.8.10) より，この値は K の値に等しい．よって K の値は **18.8.1** で与えられる．

判別分析問題に関しては (18.8.1) 式よりもむしろ行列方程式

(18.8.18)
$$|v'_{ij} - l v_{ij}| = 0$$

の固有根の方が基本的には深かろう．

しかし (18.8.18) の根と (18.8.1) の根の間には次の関係が成立している．

(18.8.19)
$$g_1 = \frac{l_1}{1+l_1}, \cdots, g_k = \frac{l_k}{1+l_k}.$$

よって (18.8.2) に変換 (18.8.19) を行なうと次の結果が得られる．

18.8.2　$\{v_{ij}\}$, $\{v'_{ij}\}$ は独立な確率変数の組で，それぞれウィシャート分布 $W(k, n, \|\sigma_{ij}\|)$, $W(k, n', \|\sigma_{ij}\|)$ を持つとする．このとき $|v'_{ij} - l v_{ij}| = 0$ の固有根 $0 < l_k < \cdots < l_1$ は次の確率素分を持つ．$0 < l_k < l_{k-1} < \cdots < l_1 < +\infty$ なる領域 E_k 上では

(18.8.20)
$$K\left[\prod_{i=1}^{k}(1+l_i)\right]^{-\frac{1}{2}(n+n')}\left[\prod_{i=1}^{k} l_i\right]^{\frac{1}{2}(n'-k-1)}\prod_{i>j=1}^{k}(l_j - l_i)\, dl_1 \cdots dl_k$$

その他では 0．ただし K は **18.8.1** で与えられる．

さて $(s+1)$ 個の標本に対する判別分析問題に戻ろう．ただし $s+1 > k$．これに関しては **18.8.2** の次の系が成立する．

18.8.2 a　$\|u^B_{ij}\|$, $\|u^W_{ij}\|$ を大きさがそれぞれ $n_1, \cdots, n_{s+1}$ の $(s+1)$ 個の標本の標本間，標本内散乱行列とする．ただし $s+1 > k$, $n_1 + \cdots + n_{s+1} - (s+1) \geqq k$．このとき

(18.8.21)
$$|u^B_{ij} - l u^W_{ij}| = 0$$

の固有根 $0 < l_k < \cdots < l_1$ は次の確率素分を持つ．

(18.8.22)
$$\begin{cases} K_0\left[\prod_{i=1}^{k}(1+l_i)\right]^{-\frac{1}{2}(n+s)}\left[\prod_{i=1}^{k} l_i\right]^{\frac{1}{2}(s-k-1)}\prod_{i>j=1}^{k}(l_j - l_i)\, dl_1 \cdots dl_k, \\ \qquad\qquad 0 < l_k < l_{k-1} < \cdots < l_1 < +\infty \\ 0, \qquad\qquad \text{その他} \end{cases}$$

ただし

(18.8.23)
$$K_0 = \pi^{\frac{1}{2}k}\prod_{i=1}^{k}\frac{\Gamma\left(\frac{n+s+1-i}{2}\right)}{\Gamma\left(\frac{n+1-i}{2}\right)\Gamma\left(\frac{s+1-i}{2}\right)\Gamma\left(\frac{k+1-i}{2}\right)}.$$

ここに $n = n_1 + \cdots + n_{s+1} - (s+1)$．

この条件の下では $\{u^B_{ij}\}$ と $\{u^W_{ij}\}$ はそれぞれウィシャート分布 $W(k, s, \|\sigma_{ij}\|)$, $W(k,$

$n, \|\sigma_{ij}\|)$ に従う独立な確率変数の組であることに注意すれば，**18.8.2** からただちに証明される．ここでも $n = n_1 + \cdots + n_{s+1} - (s+1)$．

（b）　s 個の固有値の場合 $(s < k)$

さて **18.8.1** において $n' < k$ としよう．このとき $\{v'_{ij}\}$ の要素は退化ウィシャート分布を持つであろう．事実 $\{v'_{ij}\}$ は $\left\{\sum_{\xi=1}^{n'} z_{i\xi} z_{j\xi}\right\}$ と同じ分布を持つ．ただし $(z_{1\xi}, \cdots, z_{k\xi}, \xi = 1, \cdots, n')$ は k 次元正規分布 $N(\{0\}, \|\sigma_{ij}\|)$ からの大きさ n' の標本である．よって (18.8.1) は n' 個の固有値 $0 < g_{n'} < \cdots < g_1$ を持ち $g_1, \cdots, g_{n'}$ の確率素分は (18.8.2) で k, n, n' をそれぞれ n', $n + n' - k$, k に置き換えた型に等しい．$(s+1)$ 個の標本に対する判別分析問題に立ち帰れば（ただし $s + 1 \leqq k$），(18.8.21) が固有値 $0 < l_s < \cdots < l_1$, $(s < k)$ を持つことがわかる．ゆえにこれらの固有値の確率素分は (18.8.22) において k, n, s をそれぞれ s, $n + s - k$, k に置き換えて得られる．もちろんここでも $n = n_1 + \cdots + n_{s+1} - (s+1)$．

18.8.1 および **18.8.2** で述べた結果は **18.6.2** の結果と対をなしている．これらの結果は 18.6 (b) 節のはじめに挙げた多くの研究者によって得られた．

18.9　正 準 相 関

（a）　正準相関係数の決定

本節では次の問題を考えよう．k 次元分布からの標本を仮定する．最初の s 個の成分の線形関数と最後の t 個の成分の線形関数（ただし $s + t = k$）を見つけて，この 2 つの線形関数が最大の相関係数を持つようにしたい．実際的な立場からすれば，2 つの線形関数はそれぞれ最初の s 個と最後の t 個から構成された指標とみなせよう．そして 2 つの指標のうち一方が他方を予測したり推定したりするために用いられるだろう．このように考えれば，2 つの線形関数を構成して，その相関係数をできるだけ大きくしようとすることは当然なことである．

もっと明確に述べよう．$(x_{i\xi}\,;\, i = 1, \cdots, k,\ \xi = 1, \cdots, n)$, $n > k$ を k 次元分布からの大きさ n の標本，$\|u_{ij}\|$ をこの標本の内部散乱行列とする．$(x_{p\xi}\,;\, p = 1, \cdots, s,\ \xi = 1, \cdots, n)$ を標本の最初の s 個の成分，$\|u_{pq}\|$, $p, q = 1, \cdots, s$ をこの内部散乱行列とする．

同様に $(x_{v\xi}; v=s+1,\cdots,k,\ \xi=1,\cdots,n)$, $\|u_{vw}\|$, $v,w=s+1,\cdots,k$ をそれぞれ標本の最後の t 個の成分およびその内部散乱行列とする．ただし $s+t=k$ で $t\geqq s>0$．こうすれば次のように書ける．

$$\|u_{ij}\| = \left\|\begin{array}{c:c} u_{pq} & u_{pw} \\ \hdashline u_{vq} & u_{vw} \end{array}\right\|. \tag{18.9.1}$$

ここで $\|u_{ij}\|$ は確率 1 で正則としよう．さて実ベクトル $(c_{1p}; p=1,\cdots,s)$, $(c_{2v}; v=s+1,\cdots,k)$ を考え

$$z_{1\xi} = \sum_{p=1}^{s} c_{1p}x_{p\xi}, \qquad z_{2\xi} = \sum_{v=s+1}^{k} c_{2v}x_{v\xi}, \qquad \xi=1,\cdots,n \tag{18.9.2}$$

とする．標本 $(z_{1\xi},\ z_{2\xi},\ \xi=1,\cdots,n)$ の標本平均のまわりの内部散乱行列を

$$\left\|\begin{array}{cc} \bar{u}_{11} & \bar{u}_{12} \\ \bar{u}_{21} & \bar{u}_{22} \end{array}\right\|$$

とする．問題はある基準化条件（一般性を失うことなく $\bar{u}_{11}=\bar{u}_{22}=1$ としてよい）の下で相関係数

$$R = \frac{\bar{u}_{12}}{\sqrt{\bar{u}_{11}\bar{u}_{22}}} \tag{18.9.3}$$

を最大ならしめる 2 つのベクトル $(c_{1p}; p=1,\cdots,s)$, $(c_{2v}; v=s+1,\cdots,k)$ を求めることである．

この問題に対する解は Hotelling (1935) で示されており，次のように要約される．

18.9.1　条件 $\bar{u}_{11}=\bar{u}_{22}=1$ の下で R を最大にする固有ベクトル $(c_{1p}^{(1)}; p=1,\cdots,s)$, $(c_{2v}^{(1)}; v=s+1,\cdots,k)$ は方程式系

$$\begin{aligned} -\sqrt{l_1}\sum_{q=1} u_{pq}c_{1q} + \sum_{w=s+1}^{k} u_{pw}c_{2w} &= 0, \qquad p=1,\cdots,s \\ \sum_{q=1}^{s} u_{vq}c_{1q} - \sqrt{l_1}\sum_{w=s+1}^{k} u_{vw}c_{2w} &= 0, \qquad v=s+1,\cdots,k \end{aligned} \tag{18.9.4}$$

の解である．ただし l_1 は

$$\left|\begin{array}{ccc} lu_{pq} & \cdots & -u_{pw} \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot & \cdot\ \cdot\ \cdot \\ -u_{vq} & \cdots & u_{vw} \end{array}\right| = 0 \tag{18.9.5}$$

の最大固有値である．さらに R の最大値は $\sqrt{l_1}$ になる．一般に (18.9.5) は区間 (0,1) 上の s 個の固有値 $l_1,\cdots,l_s$ を持つ．いま，標本 $(x_{i\xi}; i=1,\cdots,k; \xi=1,\cdots,n)$ はこの固有根が確率 1 ですべて異なるような分布からであると仮定する．この固有根を $0<l_s<\cdots<l_1<1$ としておこう．固有ベクトル $(c_{1p}^{(g)}; p=1,\cdots,s)$, $(c_{2v}^{(g)}; v=s+1,\cdots,k)$, $g=1,\cdots,s$ は (18.9.4) の l_1 を l_g に置き換えた方程式の根である．固有ベクトル $(c_{1p}^{(g)}; p=1,\cdots,s)$, $(c_{2v}^{(g')}; v=s+1,\cdots,k)$, $g\neq g'$ を用いて計算された R の値は 0 である．

それぞれ値 $\sqrt{l_1},\cdots,\sqrt{l_s}$ を持つ相関係数 $R^{(1)},\cdots,R^{(s)}$ はもとの標本 $(x_{1\xi},\cdots,x_{k\xi}; \xi=1,\cdots,n)$ の最初の s 個の成分 $(x_{1\xi},\cdots,x_{s\xi}; \xi=1,\cdots,n)$ と最後の t 個の成分 $(x_{s+1\xi},\cdots,x_{k\xi}; \xi=1,\cdots,n)$ との間の正準相関係数と呼ばれている．もちろん実際に最も興味ある正準相関係数は一番大きい $R^{(1)}$ である．

18.9.1 の証明は次のようにする．まず

$$\bar{u}_{11} = \sum_{p,q} u_{pq}c_{1p}c_{1q}, \qquad \bar{u}_{22} = \sum_{v,w} u_{vw}c_{2v}c_{2w}, \qquad \bar{u}_{12} = \sum_{p,w} u_{pw}c_{1p}c_{2w} \tag{18.9.6}$$

が確かめられる．ただし p, q は $1,\cdots,s$ を v, w は $s+1,\cdots,k$ を動く．さて条件 $\bar{u}_{11}=\bar{u}_{22}=1$ を満足し，R が定留関数となるベクトル $(c_{1p}; p=1,\cdots,s)$, $(c_{2v}; v=s+1,\cdots,k)$ を考察しよう．ラグランジュ乗数 λ, μ を用いれば，求めるこのベクトルは

$$\varphi = \bar{u}_{12} + (1-\bar{u}_{11})\lambda + (1-\bar{u}_{22})\mu \tag{18.9.7}$$

もまた定留関数にする．φ の極値を与えるベクトルは方程式系

$$\begin{aligned} \frac{\partial\varphi}{\partial c_{1p}} &= 0, \qquad p=1,\cdots,s \\ \frac{\partial\varphi}{\partial c_{2v}} &= 0, \qquad v=s+1,\cdots,k \end{aligned} \tag{18.9.8}$$

すなわち

$$\begin{aligned} -\lambda\sum_{q} u_{pq}c_{1q} + \sum_{w} u_{pw}c_{2w} &= 0, \qquad p=1,\cdots,s \\ \sum_{q} u_{vq}c_{1q} - \mu\sum_{w} u_{vw}c_{2w} &= 0, \qquad v=s+1,\cdots,k \end{aligned} \tag{18.9.9}$$

の解で与えられる．(18.9.9) の上の式に c_{1p} を乗じて，p に関して加え合わせ，下の式に c_{2v} を乗じて v に関して加え合わせれば，明らかに $\lambda=\mu$ になる．よって (18.9.9) の λ を μ に置き換えて，求めるベクトルについて解けばよい．しかしこの解を得るには，もちろん

(18.9.10)
$$\left|\begin{array}{c:c} -\lambda u_{pq} & u_{pw} \\ \hdashline u_{vq} & -\lambda u_{vw} \end{array}\right| = 0$$

を解かなければならない．この行列式を分解すると (18.9.10) は次のようになる．

(18.9.11)
$$(-1)^k \lambda^{k-2s} \left|\begin{array}{c:c} l u_{pq} & -u_{pw} \\ \hdashline -u_{vq} & u_{vw} \end{array}\right| = 0.$$

ただし $l = \lambda^2$．明らかに (18.9.11) の行列式，すなわち行列式 (18.9.5) は l の s 次の多項式である．この根は実数で，しかも正であることがわかる．$l_1, \cdots, l_s$ をその根とすれば，l_g に対応する固有ベクトル $(c_{1p}^{(g)}\,;\, p = 1, \cdots, s)$，$(c_{2v}^{(g)}\,;\, v = s+1, \cdots, k)$ は

(18.9.12)
$$-\sqrt{l_g} \sum_q u_{pq} c_{1q} + \sum_w u_{pw} c_{2w} = 0, \qquad p = 1, \cdots, s$$
$$\sum_q u_{vp} c_{1q} - \sqrt{l_g} \sum_w u_{vw} c_{2w} = 0, \qquad v = s+1, \cdots, k$$

の解である．この固有値を (18.9.12) に代入して，上の式を $c_{1p}^{(g)}$ 倍して p に関して加えれば $\sum_{p,q} u_{pq} c_{1p}^{(g)} c_{1q}^{(g)} = 1$ より

(18.9.13)
$$\sum_{p,w} u_{pw} c_{2w}^{(g)} c_{1p}^{(g)} = \sqrt{l_g}$$

を得る．しかし条件 $\bar{u}_{11} = \bar{u}_{22} = 1$ より (18.9.13) の左辺は (18.9.12) で与えられる 2 つの固有ベクトルを用いた (18.9.2) の z_1，z_2 の相関係数 R の値である．よって $l_g < 1$，$g = 1, \cdots, s$ で根 $l_1, \cdots, l_s$ はすべて確率 1 で異なっていると仮定しているから $(0, 1)$ 上に $l_s < \cdots < l_1$ という順序をつけておこう．l_1 は最大根だから l_1 に対応する固有ベクトルは (18.9.4) の解である．任意の標本の大きさが n $(>k)$ のとき，確率 1 で $l_1 = 1$ となるための必要十分条件は標本が抽出された k 次元分布における最初の s 個の確率変数と最後の t (ただし $s+t=k$) 個の確率変数が 1 次従属になることである．

任意の 2 つの異なる固有ベクトル $(c_{1p}^{(g)}\,;\, p = 1, \cdots, s)$ と $(c_{2v}^{(g')}\,;\, v = s+1, \cdots, k)$，$g \neq g'$ に対する z_1，z_2 の相関係数が 0 になることを示すために方程式 (18.9.12) に固有ベクトル $(c_{1p}^{(g)}\,;\, p = 1, \cdots, s)$，$(c_{2v}^{(g)}\,;\, v = s+1, \cdots, k)$ を代入して考えよう．上の式，下の式にそれぞれ $c_{1p}^{(g)}$，$c_{2v}^{(g')}$ を乗じて，それぞれ p，v に関して加える．次に g と g' を入れ換えて上と同様な操作を行なうと $l_g \neq l_{g'}$ だから方程式を適当に結び合わせれば，$\sum_{p,w} u_{pw} c_{1p}^{(g)} c_{2w}^{(g')} = 0$ になる．すなわち $(c_{1p}^{(g)}\,;\, p = 1, \cdots, s)$ と $(c_{2v}^{(g')}\,;\, v = s+1, \cdots, k)$

を用いて得られる相関係数 R は $g \neq g'$ に対して 0 になる．同様に $g \neq g'$ ならば

$$\sum_{p,q} u_{pq} c_{1p}^{(g)} c_{1q}^{(g')} = \sum_{v,w} u_{vw} c_{2v}^{(g)} c_{2w}^{(g')} = 0$$

もわかる．

(b)　正準相関係数の標本論

正準相関係数の平方 $l_1, \cdots, l_s$ の標本分布を一般的な条件の下で求める問題は大変複雑である．しかしある条件の下では $l_1, \cdots, l_s$ の分布は (18.8.2) で与えられる分布の特別な場合になる．もっと明確に述べれば次の結果になる．

18.9.2　$(x_{p\xi}\,;\, p = 1, \cdots, s,\ \xi = 1, \cdots, n)$ と $(x_{v\xi}\,;\, v = s+1, \cdots, k,\ \xi = 1, \cdots, n)$ は独立な標本で最初の標本は s 次元正規分布 $N(\{\mu_p\}\,;\, \|\sigma_{pq}\|)$ からで，次の標本は内部散乱行列 $\|u_{vw}\|$ が確率 1 で正則であるような任意の t 次元分布から抽出されたものとする．ただし $s \leqq t$，$s+t=k$．このとき (18.9.5) の根 $0 < l_s < \cdots < l_1 < 1$ は次の確率素分で与えられる分布を持つ．

(18.9.14)
$$\begin{cases} K^* \left[\prod_{p=1}^{s} (1-l_p)\right]^{\frac{1}{2}(n-t-s-2)} \left[\prod_{p=1}^{s} l_p\right]^{\frac{1}{2}(t-s-1)} \prod_{p>q=1}^{s} (l_q - l_p)\, dl_1 \cdots dl_s, \\ \qquad\qquad 0 < l_s < l_{s-1} < \cdots < l_1 < 1 \\ 0, \qquad\qquad \text{その他}. \end{cases}$$

ただし

(18.9.15)
$$K^* = \pi^{\frac{1}{2}s} \prod_{p=1}^{s} \frac{\Gamma\left(\frac{n-p}{2}\right)}{\Gamma\left(\frac{n-t-p}{2}\right)\Gamma\left(\frac{t+1-p}{2}\right)\Gamma\left(\frac{s+1-p}{2}\right)}.$$

18.9.2 の証明に際しては $(x_{v\xi}\,;\, v = s+1, \cdots, k\,;\, \xi = 1, \cdots, n)$ が任意の数で，1 次独立 (すなわち $\|u_{vw}\|$ が正則) な場合を考えれば十分である．このとき $l_1, \cdots, l_s$ の分布は (18.9.14) になることを示そう．これがわかれば (18.9.14) が任意の 1 次独立なベクトルに対して成り立つから $\{x_{v\xi}\}$ は $\|u_{vw}\|$ が確率 1 で正則な任意の t 次元分布からの標本であれば (18.9.14) が成立する．

18.9.2 の証明にすすむために，まず (18.9.5) 式の両辺に左から行列式

$$\left|\begin{array}{c:c} \delta_{pq} & \sum_v u_{pv} u^{vw} \\ \hdashline 0 & u^{vw} \end{array}\right|$$

を掛ける．(ただし; $\|u^{vw}\| = \|u_{vw}\|^{-1}$; δ_{pq} はクロネッカーの δ)．通常の行列の乗法演算を行なうと

$$(18.9.16)\qquad \left|\begin{array}{c:c} lu_{pq} - \sum\limits_{v,w} u^{vw} u_{pv} u_{wq} & -u_{pw} + \sum\limits_{v,w'} u^{vw'} u_{pv} u_{w'w} \\ \hdashline -\sum\limits_{v} u_{vq} u^{vw} & \sum\limits_{w'} u^{vw'} u_{w'w} \end{array}\right| = 0$$

を得る．ただし p, q は $1, \cdots, s$, v, w, w' は $s+1, \cdots, k$ を動く．$\sum\limits_{w'} u^{vw'} u_{w'w} = \delta_{vw}$ より (18.9.16) の右上半ブロックの各要素は 0 になる．したがって (18.9.5) の解は

$$(18.9.17)\qquad \left|\sum_{v,w} u^{vw} u_{pv} u_{wq} - lu_{pq}\right| = 0$$

の解に一致する．

さて $(x_{p\xi} - \bar{x}_p)$ を $y_{p\xi}$, $(x_{v\xi} - \bar{x}_v)$ を $y_{v\xi}$ で置き換えれば

$$(18.9.18)\qquad u_{pq} = \sum_{\xi=1}^{n} y_{p\xi} y_{q\xi} = u_{pq}^{(1)} + u_{pq}^{(2)}.$$

ただし

$$u_{pq}^{(1)} = \sum_{\xi=1}^{n}\Bigl(y_{p\xi} - \sum_{w} b_{pw} y_{w\xi}\Bigr)\Bigl(y_{q\xi} - \sum_{v} b_{qv} y_{v\xi}\Bigr)$$

$$u_{pq}^{(2)} = \sum_{v,w=s+1}^{k} u^{vw} b_{pw} b_{qv}, \qquad b_{pw} = \sum_{v=s+1}^{k} u^{vw} u_{pw}.$$

この b_{pw} を $u_{pq}^{(2)}$ に代入すれば $u_{pq}^{(2)} = \sum\limits_{v,w} u^{vw} u_{pv} u_{wq}$ となり，(18.9.17) は

$$(18.9.19)\qquad |u_{pq}^{(2)} - (u_{pq}^{(1)} + u_{pq}^{(2)}) l| = 0$$

と書ける．

18.9.2 の条件の下では確率変数の 2 つの組 $\{u_{pq}^{(1)}\}$ と $\{u_{pq}^{(2)}\}$ は独立で，それぞれウィシャート分布 $W(s, n-t-1, \|\sigma_{pq}\|)$, $W(s, t, \|\sigma_{pq}\|)$ を持つことがわかる．したがって **18.8.1** より，(18.9.17) の根は (18.8.2) の k, n', n をそれぞれ s, t, $n-t+1$ に置き換えて得られる分布を持つ．この分布は (18.9.14) である．これで **18.9.2** の証明を終える．

問　題

18.1 $(x_{1\xi}, \cdots, x_{k\xi}; \xi = 1, \cdots, n)$ が正規分布 $N(\{\mu_i\}; \|\sigma_{ij}\|)$ からの大きさ n $(>k)$

の標本ならば，$(\mu_1, \cdots, \mu_k)$ および $\|\sigma_{ij}\|$ の最尤推定量はそれぞれ $(\bar{x}_1, \cdots, \bar{x}_k)$ および $\|u_{ij}/n\|$ であることを示せ．ただし $(\bar{x}_1, \cdots, \bar{x}_k)$ は標本平均ベクトルで $\|u_{ij}\|$ は標本の内部散乱である．

18.2 対称行列 $\|a_{ij}\|$, $i, j = 1, \cdots, k < m$ の要素がウィシャート分布 $W(k, m, \|\sigma_{ij}\|)$ に従う確率変数ならば，特性関数を用いて $\|a_{pq}\|$, $p, q = 1, \cdots, s < k$ の要素がウィシャート分布 $W(s, m, \|\sigma_{pq}\|)$ を持つことを示せ．

18.3 **独立な成分を持つ k 次元正規分布からの標本相関係数の分布**　標本 $(x_{1\xi}, \cdots, x_{k\xi}; \xi = 1, \cdots, n)$, $n > k$ は k 次元正規分布 $N(\{\mu_i\}; \|\sigma_{ij}\delta_{ij}\|)$ から抽出され，$\|u_{ij}\|$ はこの標本の内部散乱行列である．このとき $r_{ij} = \dfrac{u_{ij}}{\sqrt{u_{ii}u_{jj}}}$ として相関行列 $\|r_{ij}\|$ の要素の p.d.f. は，$\|r_{ij}\|$ が正定値となる $\{r_{ij}\}$ 空間上では

$$\frac{\Gamma^k\left(\frac{n-1}{2}\right)|r_{ij}|^{\frac{1}{2}(n-k-2)}}{\pi^{k(k-1)/4}\prod\limits_{i=1}^{k}\Gamma\left(\frac{n-i}{2}\right)}$$

その他では 0 になることを示せ．

18.4 (続き)

$$\mathscr{E}(|r_{ij}|^g) = \prod_{i=2}^{k}\left[\frac{\Gamma\left(\frac{n-i}{2}+g\right)\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-i}{2}\right)\Gamma\left(\frac{n-1}{2}+g\right)}\right], \qquad g = 0, 1, 2, \cdots$$

を示せ．よって z_i をそれぞれベータ分布 $Be\left(\frac{1}{2}(n-i), \frac{1}{2}(i-1)\right)$, $i = 2, \cdots, k$ を持つ独立な確率変数とすれば，$|r_{ij}|$ の分布は $\prod\limits_{i=2}^{k} z_i$ の分布と一致することを示せ．

18.5 **標本相関係数の分布**　$\|u_{ij}\|$ は 2 次元正規分布 $N(\{\mu_i\}; \|\sigma_{ij}\|)$, $i, j = 1, 2$ からの大きさ n の標本の内部散乱行列とする．$\|u_{ij}\|$ の要素がウィシャート分布 $W(2, n-1, \|\sigma_{ij}\|)$ に従うことを利用して，標本相関係数 $r = u_{12}/\sqrt{u_{11}u_{22}}$ の p.d.f. は次の型で与えられることを示せ．

$$\begin{cases} \dfrac{(1-\rho)^{\frac{1}{2}(n-1)}(1-r^2)^{\frac{1}{2}(n-4)}}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)} \displaystyle\sum_{i=0}^{\infty} \frac{(2\rho r)^i}{i!}\Gamma^2\left(\frac{n-1+i}{2}\right), \\ \qquad\qquad\qquad -1 \leqslant r \leqslant +1 \\ 0, \qquad\qquad\quad \text{その他}. \end{cases}$$

ただし $\rho = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$ は母集団の相関係数．[Fisher (1915)]．

18.6 (続き)　ウィシャート分布 $W(2, n-1, \|\sigma_{ij}\|)$ の p.e. への変換

$$r = u_{12}/\sqrt{u_{11}u_{22}}, \qquad s = \sqrt{u_{11}u_{22}},$$

$$t = \frac{1}{2}\log(u_{22}/u_{11})$$

を行なって，r の p.d.f. は次の型でも表わされることを示せ．

$$\begin{cases} \dfrac{(n-2)\Gamma(n-1)(1-\rho^2)^{\frac{1}{2}(n-1)}(1-r^2)^{\frac{1}{2}(n-4)}(1-\rho r)^{-\frac{1}{2}(2n-3)}}{\sqrt{2}\,\pi^{3/2}} \\ \qquad \cdot \displaystyle\sum_{i=0}^{\infty} \frac{\left(\frac{1}{2}+\frac{1}{2}\rho r\right)^i \Gamma^2\left(\frac{1}{2}+i\right)}{i!\,\Gamma\left(n-\frac{1}{2}+i\right)}, & -1 \leqslant r \leqslant +1 \\ 0, & \text{その他} \end{cases}$$

[Hotelling (1953) による].

18.7 (続き)　r の確率素分をホテリング型に表わせば次のようになることを示せ.

$$\frac{(n-2)\Gamma(n-1)(1-\rho^2)^{\frac{1}{2}(n-1)}(1-r^2)^{\frac{1}{2}(n-4)}}{\sqrt{2}\,\pi\Gamma\left(n-\frac{1}{2}\right)(1-\rho r)^{\frac{1}{2}(2n-3)}}$$
$$\cdot\left\{\int_0^1\!\!\int_0^1 f(x)g(y)\left[1-\frac{1}{2}(1+\rho r)xy\right]^{-1} dy\,dx\right\} dr$$

ただし $f(x)$, $g(y)$ はそれぞれベータ分布 $Be\left(\frac{1}{2},\ n-1\right)$, $Be\left(\frac{1}{2}, \frac{1}{2}\right)$ の p.d.f. である. よって変換

$$r = \rho + \frac{1-\rho^2}{\sqrt{n}}u, \qquad x = \frac{v}{n}, \qquad y = w$$

を利用して $u = \dfrac{(r-\rho)\sqrt{n}}{1-\rho^2}$ の $n \longrightarrow \infty$ のときの極限分布は $N(0, 1)$であることを示せ.

18.8 (続き)　**9.3.1** を利用して

$$z = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right), \qquad \zeta = \frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right)$$

ならば, $n \longrightarrow \infty$ のときの $(z-\zeta)\sqrt{n}$ の極限分布は $N(0, 1)$ になることを示せ.

18.9　**正規分布の平均値ベクトルに対する信頼楕円面**　$(\bar{x}_1, \cdots, \bar{x}_k)$, $\|u_{ij}\|$ は k 次元分布 $N(\{\mu_i\};\|\sigma_{ij}\|)$ からの大きさ n $(>k)$ の標本の標本平均ベクトルおよび内部散乱行列とする.

$$n\sum_{i,j=1}^{k} u^{ij}(\mu_i-\bar{x}_i)(\mu_j-\bar{x}_j) < (1-z_\gamma)/z_\gamma$$

は母集団平均値ベクトル $(\mu_1, \cdots, \mu_k)$ に対する 100γ% 信頼領域 (楕円体) であることを示せ. ただし z_γ はベータ分布 $Be\left(\frac{1}{2}(n-k),\ \frac{1}{2}k\right)$ の右 100γ% 点である. [Hotelling (1931)].

18.10　ホテリングの T^2 と仮説《正規分布 $N(\{\mu_i\};\|\sigma_{ij}\|)$ が与えられた平均値ベクトルを持つ》に対する尤度比検定との同値性.　Ω を $\mu_1, \cdots, \mu_k$ が実数で, $\|\sigma_{ij}\|$ が正定値となる $\frac{1}{2}k(k+3)$ 次元のパラメータ空間, ω を $\mu_1 = \mu_1^\circ, \cdots, \mu_k = \mu_k^\circ$ なる Ω の部分集合であるという仮説を $\mathscr{H}(\omega\,;\Omega)$ としよう. 平均値ベクトル $(\bar{x}_1, \cdots, \bar{x}_k)$, 内部散乱行列 $\|u_{ij}\|$ を持つ $N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ からの大きさ n の標本が与えられたとき, $\mathscr{H}$ の仮説検定に対する尤度比 λ は

$$\lambda = R_1^{\frac{1}{2}n}$$

であることを示せ. ただし R_1 は (18.4.1) で与えられる. よって R_1 は (ホテリングの T^2 も) $\mathscr{H}$ の検定に対する λ と同値であり, $\mathscr{H}$ が真のとき, この分布はベータ分布 $Be\left(\frac{1}{2}(n-k),\ \frac{1}{2}k\right)$ を持つ.

18.11 (続き)　ホテリングの T^2 とマハラノビスの D^2 との仮説《等しい共分散行列を持つ 2 つの正規分布は等しい平均値ベクトルを持つ》に対する尤度比検定の同値性.　$\|u_{ij}^{(1)}\|$, $\|u_{ij}^{(2)}\|$ はそれぞれ正規分布 $N(\{\mu_i^{(1)}\}\,;\|\sigma_{ij}^{(1)}\|)$, $N(\{\mu_i^{(2)}\}\,;\|\sigma_{ij}^{(2)}\|)$ から抽出された大きさ n_1, n_2 の標本の内部散乱行列とする. ただし $n_1, n_2 > k$, $i, j = 1, \cdots, k$.

$\|u_{ij}\|$ を 2 つを合わせた標本の内部散乱行列としよう. また Ω は $\|\sigma_{ij}^{(1)}\| = \|\sigma_{ij}^{(2)}\| = \|\sigma_{ji}\|$ となるパラメータ $\{\mu_i^{(1)}\}$, $\{\mu_i^{(2)}\}$, $\|\sigma_{ij}\|$ の $\frac{1}{2}k(k+5)$ 次元空間, ω は $\mu_i^{(1)} = \mu_i^{(2)}$, $i = 1, \cdots, k$ となる Ω の $\frac{1}{2}k(k+3)$ 次元部分集合であるという仮説を $\mathscr{H}(\omega\,;\Omega)$ とする. このとき $\mathscr{H}$ の検定に対する尤度比 λ は

$$\lambda = (R_1')^{\frac{1}{2}(n_1+n_2)}$$

であることを示せ. よって (2 標本に対するホテリングの T^2, マハラノビスの D^2 ばかりでなく) R_1' も $\mathscr{H}$ の仮説検定に対する λ と同値である. $\mathscr{H}$ が真のとき, この分布はベータ分布 $Be\left(\frac{1}{2}(n_1+n_2-k-1), \frac{1}{2}k\right)$ になる.

18.12　問題 18.11 を大きさが $n_1, \cdots, n_{s+1}$ の $s+1$ 個の標本の場合へ一般化して, 仮説 $\mathscr{H}$ の検定に対する尤度比 λ は (18.5.3) で与えられる R_s と同値であることを示せ. ただし仮説 $\mathscr{H}$ は《すべて等しい共分散行列を持つ正規分布からの $(s+1)$ 個の標本が同一の平均値ベクトルを持つ》であり, (18.5.3) の $\|a_{ij}\|$ は $(s+1)$ 個の標本の内部散乱の行列の和, $\|a_{ij} + \sum_{\beta=1}^{s} b_{i\beta}b_{j\beta}\|$ は $(s+1)$ 個を 1 つに合わせた標本の内部散乱行列である. また $m = n_1 + n_2 + \cdots + n_{s+1} - (s+1)$. したがってこの場合, R_s の分布は $\mathscr{H}$ が真のとき, **18.5.1** で与えられることを示せ. ただし m は上述の値.

18.13　仮説《2 つの正規分布の共分散行列が等しい》の検定.　問題 18.11 において次の仮説 $\mathscr{H}(\omega\,;\Omega)$ を考える. Ω はそのままで, ω は $\|\sigma_{ij}^{(1)}\| = \|\sigma_{ij}^{(2)}\|$ となる Ω の部分空間とする. $\mathscr{H}$ に対する尤度比 λ は

$$\lambda = \left|\frac{u_{ij}^{(1)}}{n_1}\right|^{\frac{1}{2}n_1}\left|\frac{u_{ij}^{(2)}}{n_2}\right|^{\frac{1}{2}n_2} \Big/ \left|\frac{u_{ij}^{(1)}+u_{ij}^{(2)}}{n}\right|^{\frac{1}{2}n}$$

であることを示せ. ただし $n = n_1 + n_2$. $\mathscr{H}$ が真のとき

$$\mathscr{E}(\lambda^r) = \left[\left(\frac{n}{n_1}\right)^{n_1}\left(\frac{n}{n_2}\right)^{n_2}\right]^{\frac{1}{2}kr} \prod_{i=1}^{k} \frac{\Gamma\left(\frac{n-1-i}{2}\right)\Gamma\left(\frac{n_1-i+n_1r}{2}\right)\Gamma\left(\frac{n_2-i+n_2r}{2}\right)}{\Gamma\left(\frac{n-1-i+nr}{2}\right)\Gamma\left(\frac{n_1-i}{2}\right)\Gamma\left(\frac{n_2-i}{2}\right)}$$
$$r = 0, 1, 2, \cdots$$

を示せ.

18.14　**2 因子実験計画におけるベクトルのモデル I 分散分析**　行 $R_1, \cdots, R_r$, 列 $C_1, \cdots, C_s$ の実験配置において, ξ $(=1, \cdots, r)$ 番目の行, η $(=1, \cdots, s)$ 番目の列によ

って構成される k 次元ベクトル確率変数 $(x_{1\xi\eta},\cdots,x_{k\xi\eta})$ があるとする． rs 個のベクトル確率変数は独立で $(x_{1\xi\eta},\cdots,x_{k\xi\eta})$ は k 次元正規分布

$$N(\{\mu_i+\mu_{i\xi\cdot}+\mu_{i\cdot\eta}\},\|\sigma_{ij}\|)$$

を持つとする．ただし $\sum_{\xi}\mu_{i\xi\cdot}=\sum_{\eta}\mu_{i\cdot\eta}=0,\ i=1,\cdots,k$．さて， Ω は上述の rs 個の分布に含まれている $k(r+s-1)+\frac{1}{2}k(k+1)$ 個の独立なパラメータ空間で， ω は"行効果"がすべて 0，すなわち

$$\mu_{i1\cdot}=\mu_{i2\cdot}=\cdots=\mu_{ir\cdot}=0,\qquad i=1,\cdots,k$$

となる Ω の $(ks+\frac{1}{2}k(k+1))$ 次元部分空間であるという仮説を $\mathscr{H}(\omega\,;\Omega)$ としよう．$x_{i\xi\eta},\ \xi=1,\cdots,r,\ \eta=1,\cdots,s$ から計算した (10.6.3) の $\bar{x}_{\cdot\cdot},\ \bar{x}_{\xi\cdot},\ \bar{x}_{\cdot\eta},\ m,\ m_{\xi\cdot},\ m_{\cdot\eta}$ の値をそれぞれ $\bar{x}_{i\cdot\cdot},\ \bar{x}_{i\xi\cdot},\ \bar{x}_{i\cdot\eta},\ m_i,\ m_{i\xi\cdot},\ m_{i\cdot\eta}$ とする．

$$S_{ij\cdot\cdot}=\sum_{\xi,\eta}(x_{i\xi\eta}-m_i-m_{i\xi\cdot}-m_{i\cdot\eta})(x_{j\xi\eta}-m_j-m_{j\xi\cdot}-m_{j\cdot\eta})$$

$$S_{ij\cdot0}=\sum_{\xi,\eta}m_{i\xi\cdot}m_{j\xi\cdot}$$

とすれば $\mathscr{H}$ の検定に対する尤度比 λ は

$$\lambda=L^{\frac{1}{2}rs}$$

であることを示せ．ただし

$$L=\frac{|S_{ij\cdot\cdot}|}{|S_{ij\cdot\cdot}+S_{ij\cdot0}|}.$$

$\mathscr{H}$ が真のとき（すなわち $\mu_{i1\cdot}=\cdots=\mu_{ir\cdot}=0,\ i=1,\cdots,k$ のとき），L の分布は m,s をそれぞれ $(r-1)(s-1)$，$(r-1)$ に置き換えた **18.5.1** の R_s の分布に一致することを示せ．

18.15　最小 2 乗残差平方和の分布　$(x_{1\xi},\cdots,x_{k\xi}\,;\xi=1,\cdots,n)$，$n>k$ は k 次元分布 $N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ からの標本で， $\|u_{ij}\|$ はこの標本の内部散乱行列とする．標本を固定したとき， $\sum_{\xi=1}^{n}[x_{1\xi}-\beta_1-\beta_2x_{2\xi}-\cdots-\beta_kx_{k\xi}]^2$ の $\beta_1,\cdots,\beta_k$ に関する最小値は， $|u_{ij}|/|u_{vw}|,\ i,j=1,\cdots,k,\ v,w=2,\cdots,k$ であることを示せ．モーメントを求めるとき 18.4 節で用いた方法と同様にして

$$\mathscr{E}\left(\frac{|u_{ij}|}{|u_{vw}|}\right)^r=(2/\sigma^{11})^r\frac{\Gamma\left(\frac{n-k}{2}+r\right)}{\Gamma\left(\frac{n-k}{2}\right)},\qquad r=0,1,2,\cdots$$

を示せ． ただし σ^{11} は $\|\sigma_{ij}\|^{-1}$ の 1 行 1 列の要素． このモーメント列から， $\sigma^{11}|u_{ij}|/|u_{vw}|$ がカイ 2 乗分布 $C(n-k)$ を持つことがわかる．

18.16（続き）　重相関係数の標本分布　最小にするベクトル $(\beta_1,\cdots,\beta_k)$ の値を $(\hat{\beta}_1,\cdots,\hat{\beta}_k)$ とすれば， $x_{1\xi}$ と $\hat{\beta}_1+\hat{\beta}_2x_{2\xi}+\cdots+\hat{\beta}_kx_{k\xi},\ \xi=1,\cdots,n$ との標本相関係数 R は

$$1-R^2=\frac{|u_{ij}|}{u_{11}|u_{vw}|}$$

で与えられることを示せ．また

$$\phi(r,\theta)=\mathscr{E}\left[\left(\frac{|u_{ij}|}{|u_{vw}|}\right)^r e^{-\theta u_{11}}\right]$$

としたとき

$$\phi(r,\theta)=(1+2\sigma_{11}\theta)^{-\frac{1}{2}(n-1)}\left(\frac{1}{2}\sigma^{11}+\theta\right)^{-r}\frac{\Gamma\left(\frac{n-k}{2}+r\right)}{\Gamma\left(\frac{n-k}{2}\right)}$$

$$=\frac{\Gamma\left(\frac{n-k}{2}-r\right)}{(2\sigma_{11})^{\frac{1}{2}(n-1)}\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-k}{2}\right)}\cdot\sum_{i=0}^{\infty}\frac{\left[\frac{\sigma_{11}\sigma^{11}-1}{2\sigma_{11}}\right]^i\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{1}{2}\sigma^{11}+\theta\right)^{-\frac{1}{2}(n-1)-r-i}}{i!}$$

を示せ．

$$\mathscr{E}(1-R^2)^r=\int_0^\infty\cdots\int_0^\infty\phi(r,\xi_1+\cdots+\xi_r)\,d\xi_1\cdots d\xi_r,\qquad r=0,1,2,\cdots$$

に注意して

$$\mathscr{E}(1-R^2)^r=\frac{(1-\rho^2)^{\frac{1}{2}(n-1)}}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-k}{2}\right)}\sum_{i=0}^{\infty}\frac{\rho^{2i}\,\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(\frac{n-k}{2}+r\right)}{i!\,\Gamma\left(\frac{n-1}{2}+i+r\right)}$$

を示せ．ただし ρ^2 は $1-\rho^2=\frac{1}{\sigma_{11}\sigma^{11}}$ で与えられる母集団における x_1 と $x_2,\cdots,x_k$ との重相関係数の平方である．この $1-R^2$ の r 次のモーメントより，R^2 の p.d.f. は

$$\begin{cases}\dfrac{(1-\rho^2)^{\frac{1}{2}(n-1)}(1-R^2)^{\frac{1}{2}(n-k-2)}(R^2)^{\frac{1}{2}(k-3)}}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-k}{2}\right)}\displaystyle\sum_{i=0}^{\infty}\frac{(\rho^2R^2)^i\,\Gamma^2\left(\frac{n-1}{2}+i\right)}{i!\,\Gamma\left(\frac{k-1}{2}+i\right)}, & 0\leqslant R^2\leqslant 1\\ 0, & \text{その他}\end{cases}$$

であることを示せ．[Fisher (1928 *b*)]．

18.17（続き）　残差散乱の分布　残差

$$y_{p\xi}=\left(x_{p\xi}-\beta_p-\sum_{v=s+1}^{k}\beta_{pv}x_{v\xi}\right),\qquad p=1,\cdots,s,\quad \xi=1,\cdots,n,\quad s<k$$

とこの残差の散乱 $\left|\sum_{\xi=1}^{n}y_{p\xi}y_{q\xi}\right|$ を考える．18.4 節と同様な方法を用いて，この散乱の β の項に関する最小値は

$$|u_{ij}|/|u_{vw}|,\qquad i,j=1,\cdots,k,\quad v,w=s+1,\cdots,k$$

で，さらに標本 $(x_{1\xi},\cdots,x_{k\xi}\,;\xi=1,\cdots,n)$ が正規分布 $N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ からのとき，

$$\mathscr{E}\left(\frac{|u_{ij}|}{|u_{vw}|}\right)^r = \left(\frac{2^s|\sigma_{ij}|}{|\sigma_{vw}|}\right)^r \prod_{i=k-s+1}^{k}\left[\frac{\Gamma\left(\dfrac{n-i}{2}+r\right)}{\Gamma\left(\dfrac{n-i}{2}\right)}\right]$$

であることを示せ．よって $|u_{ij}|/|u_{vw}|$ の分布は

$$\frac{2^s|\sigma_{ij}|}{|\sigma_{vw}|}\prod_{i=1}^{s} z_i$$

の分布に等しいことを示せ．ただし $z_1,\cdots,z_s$ はそれぞれガンマ分布

$$G\left(\frac{1}{2}(n-k)\right),\ G\left(\frac{1}{2}(n-k+1)\right),\cdots,G\left(\frac{1}{2}(n-k+s-1)\right)$$

に従う独立な確率変数である．

18.18　k 次元確率分布の主成因（固有値）　共分散行列 $\|\sigma_{ij}\|$ を持つ k 次元確率分布の主成因は $\|\sigma_{ij}\|$ の固有値 $\lambda_1,\cdots,\lambda_k$（すなわち方程式 $|\sigma_{ij}-\lambda\delta_{ij}|=0$ の根）である．ただし $+\infty>\lambda_1\geqslant\cdots\geqslant\lambda_k\geqslant 0$ とする．$\lambda_1+\cdots+\lambda_k=\sigma_{11}+\cdots+\sigma_{kk}$，および $\lambda_1\lambda_2\cdots\lambda_k=|\sigma_{ij}|$ を示せ．他のすべての固有値が異なっているような固有値 λ_p に対応する長さ 1 の固有ベクトル $(c_{1p},\cdots,c_{kp})$ は

$$\sum_{j=1}^{k}(\sigma_{ij}-\lambda_p\delta_{ij})c_{jp}=0,\qquad i=1,\cdots,k$$

を満たす長さ 1 のベクトルであることを示せ．固有値 $\lambda_1,\cdots,\lambda_k$ がすべて異なっていれば，この固有値に対応する長さ 1 の固有ベクトルは互いに直交していることを示せ．よって軸を回転して新しい軸がそれぞれ固有ベクトル $(c_{11},\cdots,c_{k1}),\cdots,(c_{1k},\cdots,c_{kk})$ で与えられる方向を持つようにすれば新しい k 次元空間における確率分布は共分散行列 $\|\delta_{ij}\lambda_i\|$ を持つことを示せ．

18.19（続き）k 次元確率分布は共分散行列 $\|\sigma_{ij}\|$ を持つとする．ただし $\sigma_{ii}=\sigma^2$，$i=1,\cdots,k$，$\sigma_{ij}=\rho\sigma^2$，$i\neq j=1,\cdots,k$．このとき $\lambda_1=\sigma^2[1+(k-1)\rho]$，$\lambda_2=\cdots=\lambda_k=\sigma^2(1-\rho)$ および λ_1 に対応する長さ 1 の固有ベクトル $(c_{11},\cdots,c_{k1})$ は

$$c_{11}=\cdots=c_{k1}=\frac{1}{\sqrt{k}}$$

で与えられることを示せ．

18.20　2 組の確率変数の独立性の検定　$\|u_{ij}\|$ は k 次元分布 $N(\{\mu_i\};\|\sigma_{ij}\|)$ からの大きさ $n\ (>k)$ の標本の内部散乱行列とする．$\|u_{pq}\|$，$p,q=1,\cdots,s$，$\|u_{vw}\|$，$v,w=s+1,\cdots,k$ はそれぞれ最初の s 個，および最後の t 個 $(s+t=k,\ s\leqslant t)$ の確率変数だけを用いた標本の内部散乱行列とする．さらに Ω は $\mu_1,\cdots,\mu_k$ が実数で，$\|\sigma_{ij}\|$ が正定値となる $\dfrac{1}{2}k(k+3)$ 次元パラメータ空間で，ω は

$$\sigma_{pw}=0,\qquad p=1,\cdots,s,\quad w=s+1,\cdots,k$$

となる Ω の $\left(\dfrac{1}{2}k(k+3)-st\right)$ 次元部分集合であるという統計的仮説を $\mathscr{H}(\omega;\Omega)$ としよう．仮説検定に対するネイマン＝ピヤソンの尤度比 λ は

$$\lambda=L^{\frac{1}{2}n}$$

で与えられることを示せ．ただし

$$L=\frac{|u_{ij}|}{|u_{pq}|\cdot|u_{vw}|}.$$

18.4 (a) 節と同様にして，$\mathscr{H}$ が真のとき

$$\mathscr{E}(L^r)=\prod_{p=1}^{s}\left[\frac{\Gamma\left(\dfrac{n-p}{2}\right)\Gamma\left(\dfrac{n-t-p}{2}+r\right)}{\Gamma\left(\dfrac{n-p}{2}+r\right)\Gamma\left(\dfrac{n-t-p}{2}\right)}\right],\qquad r=0,1,2,\cdots$$

を示せ．

$s=1$，$t=k-1$ のとき，L はベータ分布 $Be\left(\dfrac{1}{2}(n-k),\ \dfrac{1}{2}(k-1)\right)$，また $s=2$，$t=k-2$ ならば $\sqrt{L}$ はベータ分布 $Be(n-k,k-2)$ を持つことを示せ．[Wilks (1935)].

18.21　正規分布の球型性の検定　k 次元正規分布 $N(\{\mu_i\};\|\sigma_{ij}\|)$ からの大きさ n の標本の内部散乱行列を $\|u_{ij}\|$ とする．さらに Ω は $\{\mu_i\}$ が実数で，$\|\sigma_{ij}\|$ が正定値となる $\dfrac{1}{2}k(k+3)$ 次元パラメータ空間で，ω は $\|\sigma_{ij}\|=\|\delta_{ij}\sigma^2\|$ となる Ω の $(k+1)$ 次元部分集合であるという仮説を $\mathscr{H}(\omega;\Omega)$ としよう．ただし δ_{ij} はクロネッカーの δ で，$\sigma^2>0$．$\mathscr{H}$ の検定に対するネイマン＝ピヤソンの尤度比 λ は

$$\lambda=L^{\frac{1}{2}n}$$

で与えられることを示せ．ただし

$$L=\frac{|u_{ij}|}{\bar{u}^k},\qquad \bar{u}=\frac{1}{k}(u_{11}+\cdots+u_{kk}).$$

このように L は $\mathscr{H}$ に対する検定としては λ と同値である，さらに $\mathscr{H}$ が真のとき

$$\mathscr{E}\left(\frac{|u_{ij}|}{\bar{u}^k}\right)^r=\int_0^\infty\cdots\int_0^\infty\phi(r,\xi_1+\cdots+\xi_{rk})\,d\xi_1\cdots d\xi_{rk}$$

$$=\frac{k^{kr}\,\Gamma\left(\dfrac{(n-1)k}{2}\right)}{\Gamma\left(\dfrac{(n-1)k}{2}+rk\right)}\prod_{i=1}^{k}\left[\frac{\Gamma\left(\dfrac{n-i}{2}+r\right)}{\Gamma\left(\dfrac{n-i}{2}\right)}\right],\qquad r=0,1,2,\cdots$$

を示せ．ただし

$$\phi(r,\theta)=\mathscr{E}(|u_{ij}|^r e^{-\theta\bar{u}})=\left(\frac{2}{\sigma^2}\right)^{kr}\left(1+\frac{2\theta}{k\sigma^2}\right)^{-\frac{1}{2}[(n-1)+2rk]}\prod_{i=1}^{k}\left[\frac{\Gamma\left(\dfrac{n-i}{2}+r\right)}{\Gamma\left(\dfrac{n-i}{2}\right)}\right].$$

$k=2$ のとき $\sqrt{L}$ はベータ分布 $Be(n-2,1)$ を持つ．[Mauchly (1940)].

18.22（続き）仮説《k 次元正規分布が k 変数に関して対称である》の検定．Ω は前問と同様で，ω は $\sigma_{ii}=\sigma^2$，$\sigma^2>0\ (i=1,\cdots,k)$，かつ $\sigma_{ij}=\sigma_{ji}=\rho\sigma^2\ (i\neq j)$，$-\dfrac{1}{k-1}\leqslant\rho<1$ を満たす Ω の部分集合であるという仮説を $\mathscr{H}^*(\omega;\Omega)$ とする．$\mathscr{H}^*$ に対するネイマン＝ピヤソンの尤度比 λ は

$$\lambda = (L^*)^{\frac{1}{2}n}$$

で与えられることを示せ．ただし

$$L^* = \frac{|u_{ij}|}{(\bar{u} - \bar{u}^*)^{k-1}(\bar{u} + (k-1)\bar{u}^*)}$$

$$\bar{u} = \frac{1}{k}(u_{11} + \cdots + u_{kk}), \qquad \bar{u}^* = \frac{1}{k(k-1)} \sum_{i \neq j} u_{ij}.$$

よって L^* は $\mathscr{H}^*$ の仮説検定に対する λ と同値である．さらに，$\mathscr{H}^*$ が真のとき

$$\mathscr{E}(L^*)^r = \int_0^\infty \cdots \int_0^\infty \psi(r, \xi_1 + \cdots + \xi_{r(k-1)}, \eta_1 + \cdots + \eta_k)\, d\xi_1 \cdots d\xi_{r(k-1)} d\eta_1 \cdots d\eta_k$$

$$= (k-1)^{r(k-1)} \frac{\Gamma\left(\frac{(n-1)(k-1)}{2}\right)}{\Gamma\left(\frac{(n-1)(k-1)}{2} + r(k-1)\right)} \prod_{i=2}^{k} \left[\frac{\Gamma\left(\frac{n-i}{2} + r\right)}{\Gamma\left(\frac{n-i}{2}\right)}\right]$$

$$r = 0, 1, 2, \cdots$$

を示せ．ただし

$$\psi(r, p, q) = \mathscr{E}\left[|u_{ij}|^r e^{-p(\bar{u} - \bar{u}^*) - q(\bar{u} + (k-1)\bar{u}^*)}\right]$$

$$= 2^{kr} \prod_{i=1}^{k} \frac{\Gamma\left(\frac{n-i}{2} + r\right)}{\Gamma\left(\frac{n-i}{2}\right)} (A^{k-1} - B)^{\frac{1}{2}(n-1)}$$

$$\cdot \left[A + \frac{2p}{k-1}\right]^{-\frac{1}{2}(k-1)[(n-1)+2r]} \cdot [B + 2q]^{-\frac{1}{2}(n-1+2r)}$$

$$A = \frac{1}{\sigma^2(1-\rho)}, \qquad B = \frac{1}{\sigma^2[1 + (k-1)\rho]}.$$

$k = 2$ のとき L^* はベータ分布 $Be\left(\frac{1}{2}(n-2), \frac{1}{2}\right)$，$k = 3$ のとき $\sqrt{L^*}$ はベータ分布 $Be(n-2, 1)$を持つことを確かめよ．[Wilks (1946)]．［検定 L^* の一般化——$N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ に対するブロック内での確率変数に関する対称性の検定——については Votaw (1948) を参照せよ］.

18.23　非心 T^2 分布　　18.4 (a) 節において標本が $N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ からならば（ただし $(\mu_1, \cdots, \mu_k) \neq (\mu_1^\circ, \cdots, \mu_k^\circ)$），(18.4.5) で定義される T^2 は p.d.f.

$$\frac{e^{-\frac{1}{2}\delta^2}(T^2)^{\frac{1}{2}k-1}[1 + T^2/(n-1)]^{-\frac{1}{2}n}}{(n-1)^{\frac{1}{2}k}\Gamma\left(\frac{n-k}{2}\right)} \cdot \sum_{i=0}^{\infty} \frac{\left[\frac{\delta^2 T^2}{2(n-1)}\right]^i \Gamma\left(\frac{n}{2} + i\right)}{i!\,\Gamma\left(\frac{k}{2} + i\right)[1 + T^2/(n-1)]^i}$$

を持つことを示せ．ただし

$$\delta^2 = n \sum_{i,j=1}^{k} \sigma_{ij}(\mu_i - \mu_i^\circ)(\mu_j - \mu_j^\circ).$$

[Hsu (1938)].

18.24　$N(\{\mu_i\}\,;\|\sigma_{ij}\|)$ からの標本の散乱 $|v_{ij}|$ のウィシャート分布を用いない母モーメント[Wilks (1934)].　$(x_{1\xi}, \cdots, x_{k\xi},\ \xi = 1, \cdots, n)$，$n \geqq k$ は k 次元正規分布 $N(\{\mu_i\}\,;$

$\|\sigma_{ij}\|)$ からの標本とする．この標本の平均 $(\mu_1, \cdots, \mu_k)$ のまわりの散乱行列を $\|v_{ij}\|$ とする．$2r < n + 1 - k$ のとき，ウィシャート分布を用いないで

$$\mathscr{E}(|v_{ij}|^{-r}) = \pi^{kr} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathscr{E}\left[\exp\left(-\sum_{p=1}^{2r} \sum_{i,j=1}^{k} v_{ij} z_{ip} z_{jq}\right)\right] \prod_{i,p} dz_{ip}$$

$$= |2\sigma_{ij}|^{-r} \prod_{i=1}^{k} \left[\frac{\Gamma\left(\frac{n+1-i}{2} - r\right)}{\Gamma\left(\frac{n+1-i}{2}\right)}\right]$$

を示せ．

参考文献

［ ］の中の数字は本文の引用頁を表わす．

R. L. Anderson (1942), Distribution of the serial correlation coefficient, *Ann. Math. Stat.*, Vol. 13, pp. 1-13. [263]

T. W. Anderson (1948), On the theory of testing serial correlation, *Skand. Aktuar.*, Vol. 31, pp. 88-116. [264]

T. W. Anderson and D. A. Darling (1952), Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes, *Ann. Math. Stat.*, Vol. 23, pp. 193-212. [166]

T. W. Anderson (1958), *An Introduction to Multivariate Statistical Analysis*, John Wiley, New York. [270]

T. W. Anderson (1960), A modification of the sequential probability ratio test to reduce the sample size, *Ann. Math. Stat.*, Vol. 31, pp. 165-197. [220]

F. C. Andrews (1954), Asymptotic behavior of some rank tests for analysis of variance, *Ann. Math. Stat.*, Vol. 25, pp. 624-736. [196]

F. J. Anscombe (1953), Sequential estimation, *Jour. Roy. Stat. Soc., Series B*, Vol. 15, pp. 1-29. [224]

H. E. Arnold (1958), *Permutation support for multivariate techniques*, Ph. D. Dissertation, Princeton University. [193]

R. R. Bahadur (1954), Sufficiency and statistical decision functions, *Ann. Math. Stat.*, Vol. 25, pp. 423-462. [83]

R. R. Bahadur (1958), Examples of inconsistency of maximum likelihood estimates, *Sankhyā*, Vol. 20, pp. 207-210. [90]

E. W. Barankin and J. Gurland (1951), On asymptotically normal efficient estimators: I, *Univ. Calif. Pub. in Stat.*, Vol. 1, No. 6, pp. 89-129. [87, 90, 108]

G. A. Barnard (1952), The frequency justification of certain sequential tests. *Biometrika*, Vol. 39, pp. 144-150. [200]

W. Bartky (1943), Multiple sampling with constant probability, *Ann. Math. Stat.*, Vol. 14, pp. 363-377. [201]

M. S. Bartlett (1950), Periodogram analysis and continuous spectra, *Biometrika*, Vol. 37, pp. 1-16. [253]

M. S. Bartlett (1953), Approximate confidence intervals, II. More than one unknown parameter, *Biometrika*, Vol. 40, pp. 306-317. [115]

M. S. Bartlett (1955), *An Introduction to Stochastic Processes*, Cambridge University Press. [242, 255]

邦訳 津村善郎・奥野忠一・築林昭明・門山允・淵脇学 訳《確率過程入門》 東京大学出版会

E. M. L. Beale (1960), Confidence regions in non-linear estimation, *Jour. Roy. Stat. Soc., Ser. B*, Vol. 22, pp. 41-88. [115]

J. Bertrand (1887), *Calcul des Probabilités*, Gauthier-Villars, Paris. [198]

G. Birkhoff and S. McLane (1953), *A Survey of Modern Algebra* (Revised Edition), Macmillan, New York. [299, 309]

邦訳 奥川光太郎・辻吉雄共訳 《現代代数学概論》改訂 3 版 白水社

Z. W. Birnbaum and F. H. Tingy (1951), One-sided confidence contours for probability distribution functions, *Ann. Math. Stat.*, Vol. 22, pp. 592-596. [61, 64]

Z. W. Birnbaum (1953 *a*), On the power of a one-sided test of fit for continuous probability functions, *Ann. Math. Stat.*, Vol. 24, pp. 484-489. [167]

Z. W. Birnbaum (1953 *b*), Distribution-free tests of fit for continuous distribution functions, *Ann. Math. Stat.*, Vol. 24, pp. 1-8. [166]

D. Blackwell (1946), On an equation of Wald, *Ann. Math. Stat.*, Vol. 17, pp. 84-87. [204]

D. Blackwell (1947), Conditional expectation and unbiased sequential estimation, *Ann. Math. Stat.*, Vol. 18, pp. 105-110. [83]

D. Blackwell and M. A. Girshick (1954), *Theory of Games and Statistical Decisions*, John Wiley, New York. [230, 240]

J. R. Blum and L. Weiss (1957), Consistency of certain two-sample tests, *Ann. Math. Stat.*, Vol. 28, pp. 242-246. [179]

M. Bôcher (1907), *Introduction to Higher Algebra*, Macmillan, New York. [11, 314]

S. Bochner (1932), *Vorlesungen über Fouriersche Integrale*, Leipzig. [248]

L. Boltzmann (1898), *Vorlesungen über Gastheorie*, Part I, J. A. Barth, Leipzig. [136]

G. E. P. Box and K. B. Wilson (1951), On the experimental attainment of optimum conditions, *Jour. Roy. Stat. Soc., Series B*, Vol. 13, pp. 1-45. [21]

G. E. P. Box (1952), Multifactor designs of first order, *Biometrika*, Vol. 39, pp. 49-57. [21]

G. E. P. Box (1954), The exploration and exploitation of response surfaces: Some general considerations and examples, *Biometrics*, Vol. 10, pp. 16-60. [21]

G. E. P. Box and J. S. Hunter (1957), Multifactor experimental designs for exploring response surfaces, *Ann. Math. Stat.*, Vol. 28, pp. 195-241. [21]

M. G. Bulmer (1957), Approximate confidence limits for components of variance, *Biometrika*, pp. 159-167. [33]

D. C. Chapman and H. Robbins (1951), Minimum variance estimation without regularity assumptions, *Ann. Math. Stat.*, Vol. 22, pp. 581-586. [78, 119]

H. Chernoff (1954), On the distribution of the likelihood ratio, *Ann. Math. Stat.*, Vol. 25, pp. 573-578. [150]

H. Chernoff and L. E. Moses (1959), *Elementary Decision Theory*, John Wiley, New York. [230]

邦訳 宮沢光一訳 現代経営科学全集 《決定理論入門》 紀の国屋書店

J. H. Chung and D. B. DeLury (1950), *Confidence Limits for the Hypergeometric Distribution*, University of Toronto Press. [96]

K. L. Chung and W. Feller (1949), Fluctuations in coin tossing, *Proc. Nat. Acad. Sci., U. S. A.*, Vol. 35, pp. 605-608. [198, 199]

C. J. Clopper and E. S. Pearson (1934), The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika*, Vol. 26, pp. 404-413. [96]

W. G. Cochran (1953), *Sampling Techniques*, John Wiley, New York. [39]

邦訳 鈴木達三・高橋宏一・脇本和昌 訳 《サンプリングの理論と方法》 全 2 巻 第 2 版 東京図書

W. G. Cochran and G. M. Cox (1957), *Experimental Designs*, second edition, John Wiley, New York. [21, 32]

D. R. Cox (1952), Sequential tests for composite hypotheses, *Proc. Camb. Phil. Soc.*, Vol. 48, pp. 290-299. [200]

H. Cramér (1928), On the composition of elementary errors, *Skand. Aktuar.*, Vol. 11, pp. 13-74, 141-180. [166]

H. Cramér (1946), *Mathematical Methods of Statistics*, Princeton University Press. [79, 80]

邦訳 池田貞雄監訳 前田功雄・松井敬 訳 《統計学の数学的方法》 全 3 巻 東京図書

J. F. Daly (1940), On the unbiased character of likelihood-ratio tests for independence in normal systems, *Ann. Math. Stat.*, Vol. 11, pp. 1-32. [134]

D. A. Darling (1957), The Kolmogorov-Smirnov, Cramér-von-Mises tests, *Ann. Math. Stat.*, Vol. 28, pp. 823-838. [67, 166]

F. N. David (1950), Two combinatorial tests of whether a sample has come from a given population, *Biometrika*, Vol. 37, pp. 97-110. [162, 163]

H. T. David and W. H. Kruskal (1956), The WAGR sequential t-test reaches a decision with probability 1, *Ann. Math. Stat.*, Vol. 27, pp. 797-805. [200]

H. T. Davis (1941), *The Analysis of Economic Time Series*, Principia Press, Bloomington, Indiana. [262]

A. P. Dempster (1955), Personal communication. [65]

A. P. Dempster (1958), A high dimensional two-sample significance test, *Ann. Math. Stat.*, Vol. 29, pp. 995-1010. [307]

W. J. Dixon (1940), A criterion for testing the hypothesis that two samples are from the same population, *Ann. Math. Stat.*, Vol. 11, pp. 199-204. [170]

W. J. Dixon (1944), Further contributions to the problem of serial correlation, *Ann.*

Math. Stat., Vol. 15, pp. 119-144. [264, 267]

W. J. Dixon and A. M. Mood (1946), The statistical sign test, *Jour. Amer. Stat. Assn.*, Vol. 41, pp. 557-566. [158]

W. J. Dixon and A. M. Mood (1948), A method of obtaining and analyzing sensitivity data, *Jour. Amer. Stat. Assn.*, Vol. 43, pp. 109-126. [201]

H. F. Dodge and H. G. Romig (1929), A method of sampling inspection, *Bell System Tech. Jour.*, Vol. 8, pp. 613-631. [124, 125, 200]

H. F. Dodge and H. G. Romig (1959), *Sampling Inspection Tables*, second edition, John Wiley, New York. [124, 201]

T. G. Donnelly (1957), *A family of sequential tests*, Ph. D. Thesis, University of North Carolina. [220]

M. D. Donsker (1952), Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems, *Ann. Math. Stat.*, Vol. 23, pp. 277-281. [67]

J. L. Doob (1949), Heuristic approach to the Kolmogorov-Smirnov theorems, *Ann. Math. Stat.*, Vol. 20, pp. 393-403. [67]

J. L. Doob (1953), *Stochastic Processes*, John Wiley, New York. [242, 266]

D. Dugué (1937), Application des proprietés de las limité au sens du calcul des probabilitiés a l'étude des diverses questions d'estimation, *Écol. Poly.*, Vol. 3, No. 4, pp. 305-372. [79, 80]

D. B. Duncan (1952), On the properties of the multiple comparison test, *Virginia Jour. of Sci.*, Vol. 3, pp. 49-67. [14]

J. Durbin and G. S. Watson (1951), Exact tests of serial correlation using noncircular statistics, *Ann. Math. Star.*, Vol. 22, pp. 446-451. [264]

A. Dvoretzky, J. Kiefer and J. Wolfowitz (1953 *a*), Sequential decision problems for processes with continuous time parameter. Testing hypotheses, *Ann. Math. Stat.*, Vol. 24, pp. 254-264. [202, 240]

A. Dvoretzky, J. Kiefer and J. Wolfowitz (1953 *b*), Sequential decision problems for processes with continuous time parameter. Problems of estimation, *Ann. Math. Stat.*, Vol. 24, pp. 403-415. [202, 240]

M. Dwass (1959), Multiple confidence procedures, *Ann. Inst. Stat. Math.*, Vol. 10, pp. 277-282. [14, 21]

C. Eisenhart (1947), The assumptions underlying the analysis of variance, *Biometrics*, Vol. 3, pp. 1-21. [33]

W. Feller (1948), On the Kolmogorov-Smirnov limit theorems for empirical distributions, *Ann. Math. Stat.*, Vol. 19, pp. 177-189. [67]

T. S. Ferguson (1958), A method of generating best asymptotically normal estimates with application to the estimation of bacterial densities, *Ann. Math. Stat.*, Vol. 29, pp. 1046-1062. [87, 91]

R. A. Fisher (1915), Frequency distribution of the values of the correlation coefficient

in samples from an indefinitely large population, *Biometrika*, Vol. 10, pp. 507-521. [323]

R. A. Fisher (1922), On the mathematical foundations of theoretical statistics, *Phil. Trans. Roy. Soc. London, Series A*, Vol. 222, pp. 309-368. [77, 79, 82, 86, 89, 107]

R. A. Fisher (1925 *a*), *Statistical Methods for Research Workers*, first edition, twelfth edition (1954), Oliver and Boyd, Edinburgh. [26]
邦訳 遠藤健児・鍋谷清治共訳 《研究者のための統計的方法》 森北出版

R. A. Fisher (1925 *b*), Theory of statistical estimation, *Proc. Camb. Phil. Soc.*, Vol. 22, pp. 700-725. [77]

R. A. Fisher (1926 *b*), On the random sequence, *Quart. Jour. Roy. Meteor. Soc.*, Vol. 52, pp. 250-258. [190, 192]

R. A. Fisher (1928 *b*), The general sampling distribution of the multiple correlation coefficient, *Proc. Roy. Soc. London, Series A*, Vol. 121, pp. 654-673. [327]

R. A. Fisher (1929), Tests of significance in harmonic analysis, *Proc. Roy. Soc. London, Series A*, Vol. 125, pp. 54-59. [258, 262]

R. A. Fisher (1930), Inverse probability, *Proc. Camb. Phil. Soc.*, Vol. 26, pp. 528-535. [96]

R. A. Fisher (1934), Two new properties of mathematical likelihood, *Proc. Roy. Soc. Series A*, Vol. 144, pp. 285-307. [120]

R. A. Fisher (1935*a*), *The Design of Experiments*, Oliver and Boyd, Edinburgh. [21]

R. A. Fisher (1935*b*), The fiducial argument in statistical inference, *Ann. Eugen.*, Vol. 6, pp. 391-398. [96, 98]

R. A. Fisher (1938), The statistical utilization of multiple measurements, *Ann. Eugen.* Vol. 8, pp. 376-386. [304]

R. A. Fisher (1939), The sampling distribution of some statistics obtained from nonlinear equations, *Ann. Eugen.*, Vol. 9, pp. 238-249. [299]

R. A. Fisher and F. Yates (1938), *Statistical Tables for Biological, Agricultural and Medical Research* (fifth edition, 1957), Oliver and Boyd, Edinburgh. [21]

D. A. S. Fraser (1957), *Nonparametric Methods in Statistics*, John Wiley, New York. [156, 192]

M. Friedman (1937), The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *Jour. Amer. Stat. Assn.*, Vol. 32, pp. 675-701. [196]

F. Garwood (1936), Fiducial limits for the Poisson distribution, *Biometrika*, Vol. 28, pp. 437-442. [96]

K. F. Gauss (1809*a*), *Theoria motus corporum coelestium in sectimibus conicis solem aurbientium*, Perthes and Besser, Hamburg. [9]

M. A. Girshick (1939), On the sampling theory of roots of determinantal equations, *Ann. Math. Stat.*, Vol. 10, pp. 203-224. [299]

B. V. Gnedenko and V. S. Koroliuk (1951), On the maximum discrepancy between two empirical distributions, *Dok. Akad. Nauk.* SSR 80, Vol. 4, pp. 525-528. [183]

F. A. Graybill (1961), *An Introduction to Linear Statistical Models*, McGraw-Hill, New York. [21, 32]

J. A. Greenwood and H. O. Hartley (1961), *Guide to Tables in Mathematical Statistics*, Princeton University Press. [21]

U. Grenander and M. Rosenblatt (1953), Statistical spectral analysis of time series arising from stationary stochastic processes, *Ann. Math. Stat.*, Vol. 24, pp. 537-558. [253]

U. Grenander and M. Rosenblatt (1957), *Statistical Analysis of Stationary Time Series*, John Wiley, New York; Almqvist and Wiksell, Stockholm. [242, 253, 255, 266]

J. B. S. Haldane (1945), On a method of estimating frequencies, *Biometrika*, Vol. 33, pp. 222-225. [229]

P. R. Halmos (1946), The theory of unbiased estimation, *Ann. Math. Stat.*, Vol. 17, pp. 34-43. [4]

P. R. Halmos and L. J. Savage (1949), Application of the Radon-Nikodym theorem to the theory of sufficient statistics, *Ann. Math. Stat.*, Vol. 20, pp. 225-241. [82]

E. J. Hannan (1955), Exact tests for serial correlation, *Biometrika*, Vol. 42, pp. 133-142. [264]

M. H. Hansen, W. N. Hurwitz and W. G. Madow (1953), *Sample Survey Methods and Theory*, Vols. I and II, John Wiley, New York. [39]

G. Herglotz (1911), Über Potenzreihen mit positivem rellen Teil im Einheits kreis, *Berichte Verh. König. Sächs. Ges. Wiss., Leipzig, Math.-Phys. Klasse*, Vol. 63, pp. 501-511. [248]

W. Hoeffding (1948*a*), A class of statistics with asymptotically normal distribution, *Ann. Math. Stat.*, Vol. 19, pp. 293-325. [190]

W. Hoeffding (1948*b*), A non-parametric test of independence, *Ann. Math. Stat.*, Vol. 19, pp. 546-557. [195]

W. Hoeffding (1952), The large-sample power of tests based on permutations of observations, *Ann. Math. Stat.*, Vol. 23, pp. 169-192. [193]

W. Hoeffding (1960), Lower bounds for the expected sample size and the average risk of a sequential procedure, *Ann. Math. Stat.*, Vol. 31, pp. 352-368. [220]

R. Hooke (1956*b*), Some applications of bipolykays to the estimation of variance components and their moments, *Ann. Math. Stat.*, Vol. 27, pp. 80-98. [32]

H. Hotelling (1931), The generalization of Student's ratio, *Ann. Math. Stat.*, Vol. 2, pp 360-378. [14, 287, 324]

H. Hotelling (1933), Analysis of a complex of statistical variables into principal components, *Jour. Educ. Psych.*, Vol. 24, pp. 417-441, 498-520. [295, 298]

H. Hotelling (1935), The most predictable criterion, *Jour. Educ. Psych.*, Vol. 26, pp. 139-142. [318]

H. Hotelling and M. R. Pabst (1936), Rank correlation and tests of significance involving no assumption of normality, *Ann. Math. Stat.*, Vol. 7, pp. 29-43. [195]

H. Hotelling (1941), Experimental determination of the maximum of a function. *Ann. Math. Stat.*, Vol. 12, pp. 20-45. [201]

H. Hotelling (1944), Some improvements in weighing and other experimental techniques, *Ann. Math. Stat.*, Vol. 15, pp. 297-306. [10]

H. Hotelling (1953), New light on the correlation coefficient and its transforms, *Jour. Roy. Stat. Soc., Series B*, Vol. 15, pp. 193-232. [324]

P. L. Hsu (1938), Notes on Hotelling's generalized T, *Ann. Math. Stat.*, Vol. 9, pp. 231-243. [330]

P. L. Hsu (1939*a*), A new proof of the joint product moment distribution, *Proc. Camb. Phil. Soc.*, Vol. 35, pp. 336-338. [282]

P. L. Hsu (1939*b*), On the distribution of roots of certain determinantal equations, *Ann. Eugen.*, Vol. 9, pp. 250-258. [299]

V. S. Huzurbazar (1948), The likelihood equation, consistency and the maxima of the likelihood function, *Ann. Eugen.*, Vol. 14, pp. 185-200. [87]

A. E. Ingham (1933), An integral that occurs in statistics, *Proc. Camb. Phil. Soc.*, Vol. 29, pp. 271-276. [282]

A. T. James (1954), Normal multivariate analysis and the orthogonal group, *Ann. Math. Stat.*, Vol. 25, pp. 40-75. [282, 299]

A. T. James (1960), The distribution of the latent roots of the covariance matrix, *Ann. Math. Stat.*, Vol. 31, pp. 151-158. [303]

G. M. Jenkins and M. B. Priestly (1957), The spectral analysis of time series, *Jour. Roy. Stat. Soc., Series B*, Vol. 19, pp. 1-12. [253]

S. Karlin and H. Rubin (1956), The theory of decision procedures for distributions with monotone likelihood ratio, *Ann. Math. Stat.*, Vol. 27, pp. 272-299. [240]

O. Kempthorne (1952), *The Design and Analysis of Experiments*, John Wiley, New York. [21, 32]

M. G. Kendall and B. B. Smith (1939), The problem of m rankings, *Ann. Math. Stat.*, Vol. 10, pp. 275-287. [196]

M. G. Kendall and R. M. Sundrum (1953), Distribution-free methods and order properties, *Rev. Int. Stat. Inst.*, Vol. 23, pp. 124-134. [156]

M. G. Kendall (1953), *Rank Correlation Methods*, second edition (first edition, 1948), Charles Griffin, London. [156]

M. G. Kendall (1957), *A Course in Multivariate Analysis*, Charles Griffin, London. [270]

J. Kiefer (1948), *Sequential Determination of the Maximum of a Function*, Ph. D. Thesis, Massachusetts Institute of Technology. [202]

J. Kiefer (1952), On minimum variance estimators, *Ann. Math. Stat.*, Vol. 23, pp. 627-

629. [78]

J. Kiefer (1953), Sequential minimax search for a maximum, *Proc. Amer. Math. Soc.*, Vol. 4, pp. 502-506. [202]

J. Kiefer (1957), Optimum sequential search and approximation methods under minimum regularity conditions, *Jour. Ind. Appl. Math.*, Vol. 5, pp. 105-136. [202]

J. Kiefer and L. Weiss (1957), Some properties of generalized sequential probability ratio test, *Ann. Math. Stat.*, Vol. 28, pp. 57-74. [220]

J. Kiefer and J. Wolfowitz (1952), Stochastic estimation of the maximum of a regression function, *Ann. Math. Stat.*, Vol. 23, pp. 462-466. [202]

B. F. Kimball (1947), Some basic theorems for developing tests of fit for the case of the non-parametric probability distribution function, I, *Ann. Math. Stat.*, Vol. 18, pp. 540-548. [166]

K. Kishen (1945), On the design of experiments for weighing and making other types of measurements, *Ann. Math. Stat.*, Vol. 16, pp. 294-300. [10]

S. Kitabatake (1958), A remark on a non-parametric test, *Math. Japan*, Vol. 4, pp. 45-49. [166]

北川・三留 (1953), *Tables for the Design of Factorial Experiments* (実験計画要因配置表), 培風館, Tokyo. [21]

A. Kolmogorov (1933*b*), Sulla determinazione empirica di una legge di distribuzione, *Giorn. dell' Inst. Ital. Attuari*, Vol. 4, pp. 83-91. [65, 67, 166]

A. Kolmogorov (1941), Stationary sequences in Hilbert space (in Russian), *Bull. Math. Univ. Moscow*, Vol. 2, No. 6. [264, 266]

B. O. Koopman (1936), On distributions admitting sufficient statistics, *Trans. Amer. Math. Soc.*, Vol. 39, pp. 399-409. [120]

T. Koopmans (1942), Serial correlation and quadratic forms in normal variables, *Ann. Math. Stat.*, Vol. 13, pp. 14-33. [264]

C. Kraft and L. LeCam (1956), A remark on the roots of the maximum likelihood equation, *Ann. Math. Stat.*, Vol. 27, pp. 1174-1177. [90]

W. H. Kruskal (1952), A nonparametric test for the several-sample problem, *Ann. Math. Stat.*, Vol. 23, pp. 525-540. [196]

S. Kullback and R. A. Leibler (1951), On information and sufficiency, *Ann. Math. Stat.*, Vol. 22, pp. 79-86. [146]

S. Kullback (1959), *Information Theory and Statistics*, John Wiley, New York. [136]

P. S. Laplace (1814), *Théorie Analytique des Probabilités*, second edition, Paris. [93]

L. LeCam (1956), On the asymptotic theory of estimation and testing hypotheses, *Proc. Third Berkeley Symp. on Math. Stat. and Prob.*, Vol. 1, pp. 129-156, University of California Press. [87]

E. L. Lehmann and C. Stein (1948), Most powerful tests of composite hypotheses. I. Normal distributions, *Ann. Math. Stat.*, Vol. 19, pp. 495-516. [130]

E. L. Lehmann and C. Stein (1949), On the theory of some non-parametric hypotheses, *Ann. Math. Stat.*, Vol. 20, pp. 28-45. [190]

E. L. Lehmann (1950), Some principles of the theory of testing hypotheses, *Ann. Math. Stat.*, Vol. 21, pp. 1-26. [130]

E. L. Lehmann and H. Scheffé (1950), Completeness, similar regions and unbiased estimates, *Sankhyā*, Vol. 10, pp. 305-340. [130]

E. L. Lehmann (1957), A theory of some multiple decision problems, *Ann. Math. Stat.*, Vol. 28, pp. 1-25. [240]

E. L. Lehmann (1959), *Testing Statistical Hypotheses*, John Wiley, New York. [121, 130, 190]

R. D. Luce and H. Raiffa (1957), *Games and Decisions*, John Wiley, New York. [230]

R. D. Luce (1959), *Individual Choice Behavior*, John Wiley, New York. [154]

W. G. Madow (1938), Contributions to the theory of multivariate statistical analysis, *Trans. Amer. Math. Soc.*, Vol. 44, pp. 454-495. [282]

P. C. Mahalanobis (1930), On tests and measures of group divergence, *Jour. Asiatic Soc. of Bengal*, Vol. 26, pp. 541-588. [291]

P. C. Mahalanobis (1936), On the generalized distance in statistics, *Proc. Nat. Inst. Sci. Calcutta*, Vol. 12, pp. 49-55. [287, 291]

P. C. Mahalanobis, R. C. Bose and S. N. Roy (1937), Normalization of statistical variates and the use of rectangular coordinates in the theory of sampling distributions, *Sankhyā*, Vol. 3, pp. 1-40. [282]

P. C. Mahalanobis (1940), A sample survey of the acreage under jute in Bengal, with discussion on planning of experiments, *Proc. Second Indian Stat. Conference*, Statistical Publishing Society, Calcutta. [201]

S. Malmquist (1954), On certain confidence contours for distribution functions, *Ann. Math. Stat.*, Vol. 25, pp. 523-542. [166]

H. B. Mann and A. Wald (1943), On the statistical treatment of linear stochastic difference equations, *Econometrica*, Vol. 11, pp. 173-220. [267]

H. B. Mann and D. R. Whitney (1947), On a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Stat.*, Vol. 18, pp. 50-60. [187, 188]

H. B. Mann (1949), *Analysis and Design of Experiments*, Dover Publications, New York. [21]

A. A. Markov (1900), *Wahrscheinlichkeitsrechumung*, Teubner, Leipzig. (German translation of Russian second edition (1908) appeared in 1912; original Russian edition appeared in 1900.) [8, 9]

F. J. Massey (1950), A note on the estimation of a distribution function by confidence limits, *Ann. Math. Stat.*, Vol. 21, pp. 116-119. [65]

F. J. Massey (1951), The distribution of the maximum deviation between two sample cumulative step functions, *Ann. Math. Stat.*, Vol. 22, pp. 125-128. [183]

H. C. Mathisen (1943), A method of testing the hypothesis that two samples are from the same population, *Ann. Math. Stat.*, Vol. 14, pp. 188-194. [170]

J. W. Mauchly (1940), Significance test for sphericity of a normal n-variate distribution, *Ann. Math. Stat.*, Vol. 11, pp. 204-209. [329]

E. J. McShane (1944), *Integration*, Princeton University Press. [73]

E. J. McShane and T. Botts (1959), *Real Analysis*, D. Van Nostrand, Princeton, N. J. [73]

R. von Mises (1931), *Wahrscheinlichskeitsrechnung, und ihre Anwendung in der Statistik und Theoretischen Physik*, Deuticke, Leipzig. [166]

A. M. Mood (1946), On Hotelling's weighing problem, *Ann. Math. Stat.*, Vol. 17, pp. 432-446. [10]

A. M. Mood (1950), *Introduction to the Theory of Statistics*, McGraw-Hill, New York. [197, 198]

A. M. Mood (1951), On the distribution of the characteristic roots of normal second-moment matrices, *Ann. Math. Stat.*, Vol. 22, pp. 266-273. [299]

A. M. Mood (1954), On the asymptotic efficiency of certain nonparametric twosample tests, *Ann. Math. Stat.*, Vol. 25, pp. 514-522. [198]

P. A. P. Moran (1948), Some theorems on time series, II. The significance of the serial correlation coefficient, *Biometrika*, Vol. 35, pp. 253-260. [264]

P. A. P. Moran, J. W. Whitfield and H. E. Daniels (1950), Symposium on ranking methods, *Jour. Roy. Stat. Soc., Series B*, Vol. 12, pp. 153-191. [156]

S. Moriguti (1954), Confidence limits for a variance component, *Reports of Statistical Applications in Research, Japanese Union of Scientists and Engineers*, Vol. 3, No. 2, pp. 29-41. [33]

R. B. Murphy (1948), Non-parametric tolerance limits, *Ann. Math. Stat.*, Vol. 19, pp. 581-589. [60]

K. R. Nair (1940), Table of confidence intervals for the median in samples from any continuous population. *Sankhyā*, Vol. 4, pp. 551-558. [56]

J. von Neumann and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, third edition 1953, Princeton University Press. [230]
邦訳 銀林浩・橋本和美・宮本敏雄監訳 《ゲームの理論と経済行動》 全 5 巻 東京図書

J. Neyman and E. S. Pearson (1928), On the use and interpretation of certain test criteria for purposes of statistical inference, *Biometrika*, Vol. 20 A, Part I, pp. 175-240, Part II, pp. 263-294. [121, 130]

J. Neyman and E. S. Pearson (1933), On the problem of the most efficient tests of statistical hypotheses, *Trans. Roy. Soc. London, Series A*, Vol. 231, pp. 289-337. [121, 125, 130]

J. Neyman (1934), On the two different aspects of the representative method: The

method of stratified samplind and the method of purposive selection, *Jour. Roy. Stat. Soc.*, Vol. 97, pp. 558-625. [42]

J. Neyman (1935), Su un teorema concernente le cosiddetti statistiche sufficienti, *Giorn. Inst. Ital. Attuari*, Vol. 6, pp. 320-334. [82]

J. Neyman (1937), Outline of a theory of statistical estimation based on the classical theory of probability, *Philos. Trans. Roy. Soc. London, Series A*, Vol. 236, pp. 333-380. [93]

M. Ogawara (1951), A note on the test of serial correlation coefficients, *Ann. Math. Stat.*, Vol, 22, pp. 115-118. [264]

M. Okamoto (1952), On a non-parametric test, *Osaka Math. Jour.*, Vol. 4, pp. 77-85. [166]

E. G. Olds (1938), Distribution of sums of squares of rank differences for small numbers of individuals, *Ann. Math. Stat.*, Vol. 9, pp. 133-148. [195]

I. Olkin (1951), *On Distribution Problems in Multivariate Analysis*, Institute of Statistics Mimeograph Series, Report No. 8, University of North Carolina. [299]

I. Olkin and S. N. Roy (1954), On multivariate distribution theory, *Ann. Math. Stat.*, Vol. 25, pp. 329-339. [299]

E. Parzen (1957), On consistent estimates of the spectrum of a stationary time series, *Ann. Math. Stat.*, Vol. 28, pp. 329-348. [253]

K. Pearson (1900), On a criterion that a system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling, *Phil. Mag.*, Vol. 50, pp. 157-175. [115]

E. J. G. Pitman (1936), Sufficient statistics and intrinsic accuracy, *Proc. Camb. Phil. Soc.*, Vol. 32, pp. 567-579. [120]

E. J. G. Pitman (1937*b*), Significance tests which may be applied to samples from any populations, *Jour. Roy. Stat. Soc., Series B*, Vol. 4, pp. 119-130. [192]

E. J. G. Pitman (1938), Significance tests which may be applied to samples from any populations, III. The analysis of variance test, *Biometrika*, Vol. 29, pp. 322-335. [193]

M. H. Quenouille (1948), Some results in the testing of serial correlation coefficients, *Biometrika*, Vol. 35, pp. 261-267. [256, 264]

H. Raiffa and R. Schlaifer (1961), *Applied Statistical Decision Theory*, Graduate School of Business Administration, Harvard University. [230]

C. R. Rao (1945), Information and accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.*, Vol. 37, pp. 81-91. [79, 80, 83]

C. R. Rao (1949), Sufficient statistics and minimum variance estimates, *Proc. Camb. Phil. Soc.*, Vol. 45, pp. 213-218. [90, 120]

C. R. Rao (1952), *Advanced Statistical Methods in Biometric Research*, John Wiley, New York. [270]

G. Rasch (1948), A functional equation for Wishart's distribution, *Ann. Math. Stat.*, Vol. 19, pp. 262-266. [282]

W. E. Ricker (1937), The concept of confidence or fiducial limits applied to the Poisson frequency distribution, *Jour. Amer. Stat. Assn.*, Vol. 32, pp. 349-356. [96]

H. Robbins (1944*b*), On distribution-free tolerance limits in random sampling, *Ann. Math. Stat.*, Vol. 15, pp. 214-216. [60]

H. Robbins (1952), Some aspects of the sequential design of experiments, *Bull. Amer. Math. Soc.*, Vol. 58, pp. 527-535. [202]

H. Robbins and S. Monro (1951), A stochastic approximation method, *Am. Math. Stat.*, Vol. 22, pp. 400-407. [202]

S. N. Roy (1939), *p*-statistics or some generalizatios in analysis of variance appropriate to multivariate problems, *Sanhkyā*, Vol. 4, pp. 381-396. [299]

S. N. Roy and R. C. Bose (1953), Simultaneous confidence interval estimation, *Ann. Math. Stat.*, Vol. 24, pp. 513-536. [14]

S. N. Roy (1954), Some further results in simultaneous confidence interval estimation, *Ann. Math. Stat.*, Vol. 25, pp. 752-761. [14]

S. N. Roy (1957), *Some Aspects of Multivariate Analysis*, John Wiley, New York; Indian Statistical Institute, Calcutta. [270]

S. Rushton (1950), On a sequential t-test, *Biometrika*, Vol. 37, pp. 326-333. [200]

S. Rushton (1952), On a two-sided sequential t-test, *Biometrika*, Vol. 39, pp. 302-308. [200]

S. Saks (1937), *Theory of the Integral*, Second Edition (English translation by L. C. Young), Stechert, New York. [73]

I. R. Savage (1962), *Bibliography of nonparametric statistics*, Harvard University Press. Cambridge. [156]

L. J. Savage (1954), *The Foundations of Statistics*, John Wiley, New York. [230]

S. R. Savur (1937), The use of the median in tests of significance, *Proc. Indian Acad. Sci.*, Section A, Vol. 5, pp. 564-576. [56]

H. Scheffé (1943), Statistical inference in the non-parametric case, *Ann. Math. Stat.*, Vol. 14, pp. 305-332. [156]

H. Scheffé (1953), A method for judging all contrasts in the analysis of variance, *Biometrika*, Vol. 40, pp. 87-104. [14, 15]

H. Scheffé (1959), *The Analysis of Variance*, John Wiley, New York. [21, 32, 33]

A. Schuster (1898), On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena, *Terrestrial Magnetism*, Vol. 3, pp. 13-41. [258]

W. F. Sheppard (1898), On the calculation of the most probable values of frequency-constants for data arranged according to equidistant divisions of a scale, *Proc. London Math. Soc.*, Vol. 29, pp: 353-380. [52]

C. E. Shannon (1948), A mathematical theory of communication, *Bell System Tech. Jour.*, Vol. 27, pp. 379-423 and 623-656. [137]

B. Sherman (1950), A random variable related to the spacing of sample values, *Ann. Math. Stat.*, Vol. 21, pp. 339-361. [166]

W. A. Shewhart (1931), *Economic Control of Quality of Manufactured Product*, D. Van Nostrand, Princeton, N. J. [59]

E. Slutzky (1937), The summation of random causes as the source of cyclic processes, *Econometrica*, Vol. 5, pp. 105-146. [267]

N. Smirnov (1939*a*), On the estimation of the discrepancy between empirical curves of distribution for two independent samples, *Bull. Math. Univ. Moscow*, Vol. 2, No. 2, pp. 3-16. [61, 65, 170, 182]

N. Smirnov (1939*b*), Sur les écarts de la courbe de distribution empirique (In Russian, with French summary), *Rec. Math. N. S.*, Vol. 6, pp. 3-26. [166]

N. Smirnov (1944), Approximation of a distribution function from a sample (in Russian), *Uspehi Matemat. Nauk*, Vol. 10, pp. 179-206. [61]

P. N. Somerville (1958), Tables for obtaining non-parametric tolerance limits, *Am. Math. Stat.*, Vol. 29, pp. 599-601. [60]

Statistical Research Group, Columbia University (1945), *Sequential Analysis of Statistical Data: Applications*, Columbia University Press. [224]

Statistical Research Group, Columbia University (1947), *Techniques of Statistical Analysis*, McGraw-Hill, New York. [224]

C. Stein (1945), A two-sample test for a linear hypothesis whose power is independent of the variance, *Ann. Math. Stat.*, Vol. 16, pp. 243-258. [225]

C. Stein (1946), A note on cumulative sums, *Ann. Math. Stat.*, Vol. 17, pp. 498-499. [212]

C. Stein and A. Wald (1947), Sequential confidence intervals for the mean of a normal distribution with known variance, *Ann. Math. Stat.*, Vol. 18, pp. 427-433. [224]

F. F. Stephan and P. J. McCarthy (1958), *Sampling Opinion*, John Wiley, New York. [39]

P. V. Sukhatme (1954), *Sampling Theory of Surveys with Applications*, Iowa State College Press, Ames, Iowa. [39]

E. Sverdrup (1947), Derivation of the Wishart distribution of the second order sample moments by straightforward integration of a multiple intergral, *Skand. Aktuar*, Vol. 30, pp. 151-166. [282]

W. R. Thompson (1936), On confidence ranges for the median and other expectation distributions for populations of unknown distribution form, *Ann. Math. Stat.*, Vol. 7, pp. 122-128. [56]

G. Tintner (1940), *The Variate Difference Method*, Principia Press, Blooming-ton, Indiana. [256]

J. W. Tukey (1949*a*), The sampling theory of power spectrum estimates, *Proceedings on Applications of Autocorrelation Analysis to Physical Problems* (NAVEXOS-P-735). [253, 254]

J. W. Tukey (1953), *The Problem of Multiple Comparisons*, Unpublished Manuscript, Princeton University. [14, 15, 18]

J. W. Tukey (1956*b*), Variances of variance components: I. Balanced designs, *Ann. Math. Stat.*, Vol. 27, pp. 722-736. [32]

J. W. Tukey (1957*a*), Variances of variance components: II. The unbalanced single classification, *Ann. Math. Stat.*, Vol. 28, pp. 43-56. [32]

J. W. Tukey (1957*b*), Variances of variance components: III. Third moments in a balanced single classification, *Ann. Math. Stat.*, Vol. 28, pp. 378-384. [32]

J. W. Tukey (1957*c*), Some examples with fiducial relevance, *Ann. Math. Stat.*, Vol. 28, pp. 687-695. [98]

D. F. Votaw, Jr. (1948), Testing compound symmetry in a normal multivariate distribution, *Ann. Math. Stat.*, Vol. 19, pp. 447-473. [330]

A. Wald and J. Wolfowitz (1939), Confidence limits for continuous distribution functions, *Ann. Math. Stat.*, Vol. 10, pp. 105-118. [65]

A. Wald (1939), Contributions to the theory of statistical estimation and testing hypotheses, *Ann. Math. Stat.*, Vol. 10, pp. 299-326. [128, 129]

A. Wald and J. Wolfowitz (1940), On a test whether two samples are from the same population, *Ann. Math. Stat.*, Vol. 11, pp. 147-162. [123, 170, 180, 182]

A. Wald (1941*a*), Asymptotically most powerful tests of statistical hypotheses, *Ann. Math. Stat.*, Vol. 12, pp. 1-19. [130, 142]

A. Wald (1941*b*), Some examples of asymptotically most powerful tests, *Ann. Math. Stat.*, Vol. 12, pp. 396-408. [130]

A. Wald (1942), Asymptotically shortest confidence intervals, *Ann. Math. Stat.*, Vol. 13, pp. 127-137. [102]

A. Wald and J. Wolfowitz (1944), Statistical tests based on permutations of the observations, *Ann. Math. Stat.*, Vol. 15, pp. 358-372. [192]

A. Wald (1945), Sequential tests of statistical hypotheses, *Ann. Math. Stat.*, Vol. 16, pp. 117-186. [201, 204, 211, 215, 218, 220, 222]

A. Wald and J. Wolfowitz (1946), Tolerance limits for a normal distribution, *Ann. Math. Stat.*, Vol. 17, pp. 208-215. [69]

A. Wald (1947*a*), *Sequential Analysis*, John Wiley, New York. [200, 202, 217, 218, 224]

A. Wald (1947*b*), Foundations of a general theory of statistical decision functions, *Econometrica*, Vol. 15, pp. 279-313. [230]

A. Wald (1948), Asymptotic properties of the maximum likelihood estimate of an unknown parameter of a discrete stochastic process, *Ann. Math. Stat.*, Vol. 19, pp. 40-46. [87, 90]

A. Wald and J. Wolfowitz (1948), Optimum character of the sequential probability ratio test, *Ann. Math. Stat.*, Vol. 19, pp. 326-339. [220]

A. Wald (1949*a*), Statistical decision functions, *Ann. Math. Stat.*, Vol. 20, pp. 165-205. [230]

A. Wald (1949*b*), Note on the consistency of the maximum likelihood estimate, *Ann. Math. Stat.*, Vol. 20, pp. 595-601. [86, 87]

A. Wald (1950), *Statistical Decision Functions*, John Wiley, New York. [230, 233, 240]

W. A. Wallis (1939), The correlation ratio for ranked data, *Jour. Amer. Stat. Assn.*, Vol. 34, pp. 533-538. [196]

J. E. Walsh (1949), Some significance tests for the median which are valid under very general conditions, *Ann. Math. Stat.*, Vol. 20, pp. 64-81. [158]

L. Weiss (1961), *Statistical Decision Theory*, McGraw-Hill, New York. [230]

B. L. Welch (1937), On the z-test in randomized blocks and Latin squares, *Biometrika*, Vol. 29, pp. 21-52. [193]

P. Whittle (1951), *Hypothesis Testing in Time Series*, University of Uppsala. [264]

N. Wiener (1949), *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, John Wiley, New York. [264, 266]

F. Wilcoxon (1945), Individual comparisons by ranking methods, *Biometrics*, Vol. 1, pp. 80-83. [187]

S. S. Wilks (1932), Certain generalizations in the analysis of variance, *Biometrika*, Vol. 24, pp. 471-494. [284, 294]

S. S. Wilks (1934), Moment-generating operators for determinants of product moments in samples from a normal system, *Ann. Math.*, Vol. 35, pp. 312-340. [330]

S. S. Wilks (1935), On the independence of k sets of normally distributed statistical variables, *Econometrica*, Vol. 3, pp. 309-326. [329]

S. S. Wilks (1938*a*), The large-sample distribution of the likelihood ratio for testing composite hypotheses, *Ann. Math. Stat.*, Vol. 9, pp. 60-62. [130, 147]

S. S. Wilks (1938*b*), Shortest average confidence intervals from large samples, *Ann. Math. Stat.*, Vol. 9, pp. 166-175. [102]

S. S. Wilks and J. F. Daly (1939), An optimum property of confidence regions associated with the likelihood function, *Ann. Math. Stat.*, Vol. 10, pp. 225-235. [115]

S. S. Wilks (1941), On the determination of sample sizes for setting tolerance limits, *Ann. Math. Stat.*, Vol. 12, pp. 91-96. [59, 69]

S. S. Wilks (1942), Statistical prediction with special reference to the problem of, tolerance limits, *Ann. Math. Stat.*, Vol. 13, pp. 400-409. [59]

S. S. Wilks (1946), Sample criteria for testing equality of means, equality of variances and equality of covariances in a normal multivariate distribution, *Ann. Math. Stat.*, Vol. 17, pp. 257-281. [330]

S. S. Wilks (1948), Order statistics, *Bull. Am. Math. Soc.*, Vol. 54, pp. 5-50. [156]

S. S. Wilks (1959*a*), Nonparametric statistical inference, *Probability and Statistics, The Harald Cramér Volume*, Almqvist and Wiksell, Stockholm. [156]

S. S. Wilks (1960*a*), Multidimensional statistical scatter, *Contributions to Probability and Statistics in Honor of Harold Hotelling*, Stanford University Press. [277]

S. S. Wilks (1960*b*), A two-stage scheme for sampling without replacement, *Bull. de l'Institut Intern.*, Vol. 37, No. 2, pp. 241-248. [44]

S. S. Wilks (1961), A combinatorial test for the problem of two samples from continuous distributions, *Proc. Fourth Berkeley Symp. on Math. Stat. and Prob.*, University of California Press. [170, 179]

E. B. Wilson (1927), Probable inference, the law of succession, and statistical inference, *Jour. Amer. Stat. Assn.*, Vol. 27, pp. 209-212. [93]

J. Wishart (1928), The generalized product moment distribution in samples from a normal multivariate population, *Biometrika*, Vol. 20A, pp. 32-52. [277, 281, 282]

J. Wishart and M. S. Bartlett (1932), The distribution of the second order moment statistics in a normal system, *Proc. Camb. Phil. Soc.*, Vol. 28, pp. 455-459. [282]

H. Wold (1938), *A Study in the Analysis of Stationary Times Series*, University of Uppsala. [242, 267]

J. Wolfowitz (1949), Non-parametric statistical inference, *Proc. (First) Berkeley Symposium on Math, Stat. and Prob.*, pp. 93-113, University of California Press. [156]

H. Working and H. Hotelling (1929), Application of the theory of error to the interpretation of trends, *Jour. Amer. Stat. Assn.*, March Supplement, pp. 73-85. [21]

R. Wormleighton (1959), Some tests of permutaion symmetry, *Ann. Math. Stat.*, Vol. 30, pp. 1005-1017. [196]

F. Yates (1949), *Sampling Methods for Censuses and Surveys*, Charles Griffin, London. [39]

索　　引

ア　行

カ　行

訳者あとがき

本書は SAMUEL S. WILKS 著 *MATHEMATICAL STATISTICS*, 1962 年増訂新版の全訳である．翻訳にあたっては原著第 1 章から第 9 章までを第 1 分冊，第 10 章から第 18 章までを第 2 分冊とした．第 1 章から第 4 章まで，第 10 章から第 13 章までを田中英之，第 5 章から第 9 章まで，第 14 章から第 18 章までを岩本誠一が分担した．

コルモゴロフにより近代確率論が構築されて以来，数理統計学は近年急速に発展して来ている．本書は初版が 1943 年に出版されているが，1962 年のこの増訂新版では新しくノンパラメトリック推定，ノンパラメトリック検定，統計的逐次解析，統計的決定関数論，時系列等が追加されている．特に，ノンパラメトリック手法は今日多くの研究者のテーマでもあり，逐次解析，決定関数論は多段決定過程，ゲームの理論等の統計的制御過程へと発展しつつある．また時系列も確率過程自身として確率論の主題になっている．このような統計学の流れを適確にとらえて，初版を大幅に組み替えている．この意味で，本書は数理統計学の名著にふさわしい内容を整えているといえる．

翻訳にあたっては原著に忠実なことを旨としたが，明らかに誤りと思われる点については訂正した．しかし誤訳や訳出不足の点があれば，それは訳者の責任であり，読者の御指摘を受け補正して行きたい．

最後に浅学菲才な我々に翻訳の機会を与えてくださった東京図書に心から感謝の意を表し，合わせて我々のわがままな訳出をいつも笑顔で処理していただいた編集部の桐村恭子氏をはじめ皆様に謝辞を贈りたい．

1972 年 11 月　　訳　者

訳者紹介

田中英之　1962年　東京教育大学理学部数学科卒業
　　　　　現　在　味の素株式会社中央研究所

岩本誠一　1970年　九州大学大学院理学研究科数学
　　　　　　　　　専攻修士課程修了
　　　　　　　　　福岡大学理学部講師を経て
　　　　　現　在　九州大学理学部助手

数理統計学 2

1972年11月16日　第1刷発行

検印省略

著者　S. S. ウィルクス

訳者　田中英之
　　　岩本誠一

発行所　東京図書株式会社

本社　東京都新宿区本塩町23（タナカビル）
振替東京 13803 電話 (357) 0636〜7
営業所　東京都文京区水道1－8－1
電話 (814) 7818〜9

落丁乱丁はお取替えします

3341-2057-5160